2309

REGULATION OF EXPRESSION OF THE BACTERIOPHAGE M13 GENOME

Initiation and termination of transcription

Promotor:
Prof.dr. J.G.G. Schoenmakers
Co-referent:
Dr. R.N.H. Konings

REGULATION OF EXPRESSION OF THE BACTERIOPHAGE M13 GENOME

Initiation and termination of transcription

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE WISKUNDE EN NATUURWETENSCHAPPEN AAN DE KATHOLIEKE UNIVERSITEIT TE NIJMEGEN, OP GEZAG VAN DE RECTOR MAGNIFICUS PROF.DR. A.J.H. VENDRIK, VOLGENS BESLUIT VAN HET COLLEGE VAN DECANEN IN HET OPENBAAR TE VERDEDIGEN OP VRIJDAG 15 DECEMBER 1978 DES NAMIDDAGS OM 2.00 UUR PRECIES

DOOR

LUPPO EDENS

GEBOREN TE GRONINGEN

1978

DRUK: STICHTING STUDENTENPERS NIJMEGEN

Allen die hebben bijgedragen aan het tot stand komen van dit proefschrift betuig ik gaarne mijn erkentelijkheid.

In de eerste plaats denk ik hierbij aan Peter van Wezenbeek en aan Guus Simons, die een wezenlijke bijdrage aan het onderzoek hebben geleverd. Verder wil ik Josephine Jansen noemen, die altijd op een bijzonder betrouwbare manier de grote reeksen translatie-experimenten heeft uitgevoerd en Gonda van Duren, die op voortvarende wijze diverse restrictie-enzymen heeft geïsoleerd.

Verder ben ik veel dank verschuldigd aan Dr. R. Schilperoort en aan Drs. R. van Meteren van de Universiteit van Leiden. De hoge kwaliteit van de door hen verstrekte RNA polymerase-preparaten is ongetwijfeld een belangrijke factor geweest voor het welslagen van de in dit proefschrift beschreven experimenten.

Tenslotte wil ik Dr. D. Waddington bedanken voor het corrigeren van het Engels van de diverse hoofdstukken en de Directie van Unilever Research voor de bijzonder genereuze wijze waarop zij mij in de gelegenheid heeft gesteld om mijn proefschrift te voltooien.

CONTENTS

Part of the experimental work described in this thesis has already been published or has been accepted for publication.

CHAPTER III - Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Physical mapping of the central terminator for transcription on the bacteriophage M13 genome. *Nucleic Acid Res.*, 2, 1811-1820.

CHAPTER IV - Edens, L., van Wezenbeek, P., Konings, R.N.H. and Schoenmakers, J.G.G. (1976). Mapping of promoter sites on the genome of bacteriophage M13. *Eur. J. Biochem.*, 70, 577-587.

CHAPTER V - Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). A cascade mechanism of transcription in bacteriophage M13 DNA. *Virology*, 86, 354-367.

CHAPTER VII - Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). Transcription of bacteriophage M13 DNA existence of promoters directly preceding genes III, VI and I. *J. Virol.*, accepted for publication.

CHAPTER I

INTRODUCTION

1.1. General introduction

The single-stranded DNA phages fall into two general classes which are denoted 'filamentous' and 'icosahedral' on the basis of the shape of their representants. Although these two phage classes show several similarities with respect to the size and the organisation of their genomes, the life cycles of the filamentous and icosahedral phages are quite different. The icosahedral phages, to which class belong for instance the phages ØX-174, G4 and S13, follow the conventional infective cycle of adsorption, reproduction and release of progeny particles accomplished by lysis of the host. In contrast, virion reproduction of the filamentous phages is not accompanied by killing of the host cell. In this case the infected bacteria continue to grow and to divide while progeny virions are extruded continuously through the cell membrane. Another difference between the filamentous and the icosahedral phages is the specificity exhibited with regard to the host cell. In contrast to the icosahedral phages, the filamentous phages require host cells which contain sex pili. These sex pili are long filaments which are present in several copies per cell and which are specified by episomal DNA structures within the bacterial cell. Apart from their role in the bacterial conjugation event, sex pili function also as a specific adsorption site for filamentous phages. On the basis of their molecular architecture the filamentous phages can be divided into two distinct classes. The class I filamentous phages (1) includes both F-pilus and I-pilus specific organisms which are all hosted by Escheria coli. The F-pilus specific (Ff) phages are represented by the phages fd, f1, M13, ZJ2, Ec9, AE2, HR and δA (2) and the I-pilus specific (If) phages by If1, If2 (3) and IKe (4) The class II filamentous phages, which differ from the class I phages with regard to their X-ray diffraction patterns (5) included the phages Pf and Xf which are hosted by Pseudomonas aeruginosa (6) and by Xanthomonas oryzae (7) respectively.

Of the filamentous phages, only the group of the Ff phages is both genetically and biochemically well characterized. Essentially all of the understanding of this particular group of pha-

ges has been attained using only a few members of the Ff group, i.e. the bacteriophages f1, fd and M13. These three rod-shaped phages were described for the first time in the early sixties (8, 9, 10) and since that time they have been the subject of very intensive studies on the level of DNA replication, transcription and protein synthesis (11, 12).In this thesis special attention has been paid to the gene expression mechanism of the bacteriophage M13 genome. For that purpose, the transcription mechanism of the phage M13 replicative form DNA was studied in detail with the aid of an *in vitro* RNA synthesising system and an *in vitro* protein synthesising system.

1.2. Structure of the Ff virion

The Ff virion is a rod-shaped structure of about 850 nm long and 6 nm wide which consists of a protein coat surrounding the viral DNA molecule (2). In the protein coat the molecules overlap each other like shingles forming a cylindrical shell in which the single-stranded nucleic acid (molecular weight 2×10^6) is ensheathed as an elongated loop structure (1,10,13,14,15). The virion coat is composed of two proteins. The major coat protein ('B'-protein) is present in 2000-3000 copies and accounts for virtually all of the protein mass of the virion coat. The amino acid sequence of this protein has been determined for phage fd (16,17), whereas for several other members of the Ff group the overall amino acid composition of this protein has been determined (16,18,19, 20). According to these sequence data the major coat protein consists of 50 amino acids and has a molecular weight of 5200.

The minor coat protein of the virion ('A'-protein) is present in only a few copies at one tip of the filament (21, 22, 23, 24). Estimates of the size of this protein differ greatly and range from 42,000 (28, Van Wezenbeek *et al.*, in preparation) up to 68,000 (26). Mutant phage particles defective in gene III are abnormal in the process in which the infecting virion is adsorbed to the host cell and it is, therefore, presumed that the minor coat protein plays an important role in the phage adsorption process or in the immediately following penetration event (27).

Apart from the proteins A and B, the coat of the Ff phages possibly contains a third protein product. This additional compo-

nent, also called 'C'-protein, has a molecular weight of about 2.000 and is present in 20-30 copies (26). Although the genetic origin of this 'C'-protein is unknown, recent evidence suggests that it represents an additional viral gene product (28, Simons et al., in preparation).

1.3. Life cycle of the Ff virion

Based upon electron-microscopic observations, it is generally accepted that in the initial phase of infection the Ff phages are attached to the tip of the F-pilus (2). The adsorbed phages are then transported to the cell surface, possibly as a result of retraction of the F-pilus (2, 29) and here the phage disintegrates. In the disintegration process, which is possibly coupled with the initiation of DNA replication (30), the major coat protein (the 'B' or 'gene VIII' protein) is stripped of the DNA and deposited in the inner cell membrane until reused for the synthesis of progeny phage particles (31, 32). In contrast, the minor coat protein (the 'A' or 'gene III' protein) remains in close association with the viral DNA (27). Some observations indicate that this protein functions analogously to the cistron H protein of the icosahedral phage ØX-174 (33), as a 'pilot' protein in the replication process in which the viral DNA is multiplied (30). If true, the gene III protein promotes the interaction of the viral DNA with one of the membrane-bound replication systems of the host.

The consecutive steps in the Ff-DNA synthetic process involve:

1. replication of the penetrating single-stranded DNA yielding a double-stranded, replicative form DNA molecule (SS → RF).
2. multiplication of the RF molecule yielding a pool of RF molecules (RF → RF).
3. production of progeny single-stranded DNA by displacement of the viral strand from the RF molecule and concomitant synthesis of new viral stands (RF → SS).

The first step in this process, the conversion of the viral DNA in a duplex replicative form molecule, requires the synthesis of a complementary DNA strand. This reaction is completely dependent upon the DNA replication system of the host and, in addition, upon the host's RNA polymerase (34, 35, 16, 36,37). Based upon these

observations and the results of *in vitro* experiments performed by Kornberg's group, a model for the conversion of the viral DNA to RF was proposed (37, 38). According to this model, synthesis of the complementary DNA strand is initiated by the synthesis of a short RNA primer (39). This primer molecule is generated in a double-stranded loop structure in the viral DNA and, subsequently, extended by DNA polymerase III holoenzyme. After a complete round of synthesis, the primer RNA molecule is removed and the resulting gap in the complementary DNA strand (40) is closed by DNA polymerase I in combination with polynucleotide ligase. By analogy with recent work on phage ØX-174, it is most likely that the double-stranded, covalently closed and relaxed DNA structure so generated, is converted in a next step into the supertwisted RF-I molecule by the action of DNA gyrase (41, 42).

During the initial 10 minutes of the infection cycle, the rate of DNA synthesis in infected cells exceeds that of uninfected cells by severalfold (43). This stimulation of DNA synthesis is due to the synthesis of some 100-200 RF molecules in a process in which the parental RF-I molecule is replicated. In the complex RF replication process several bacterial functions are involved (34, 35, 44, 45, 46, 47, 48), but only a single viral gene product; the gene II protein (49). Upon infection with phage particles defective in gene II, the parental RF is formed although the replication of this molecule is inhibited. As suggested by the observation that in these experiments reduced amounts of RF II molecules were also observed (RF molecules with a nick in one DNA strand; 50, 51, 35), it has been proposed that the gene II protein functions as a strand-specific endonuclease involved in generating RF II molecules. Since RF II molecules probably act as immediate precursors in the replication mechanism leading to the formation of the pool of RF molecules, initiation of the RF replication is thought to occur at the 3'-OH terminus generated in the viral DNA strand by the gene II product. The site on the phage genome where the single-strand nick is introduced, is almost certainly located in the same region where the RNA primer for complementary DNA strand synthesis is also made, that is in the boundary of gene II and gene IV (40, 52, 53). However, the exact position of this site has not yet been determined.

Although, according to the rate of DNA synthesis, RF replication terminates very soon after infection (43), there is no sharp division between the period in which RF replication occurs and the period of single-strand synthesis. This is due in all probability to the fact that the switch-over between the two processes occurs gradually. As shown by Mazur and Model (54) and by Mazur and Zinder (55), the balance between the two stadia is regulated by the amount of free gene V protein. Due to the expression of increasing numbers of RF molecules, increasing amounts of this protein become available and when this amount has been built up to a critical level (about 10^5 molecules/cell; 56) the RF replication process is switched over to single-strand synthesis (57). The regulatory role of gene V protein in this process is based on its capacity to bind strongly and cooperatively to single-stranded DNA (58, 56). The viral DNA strands formed in the RF replication process are quickly covered with gene V protein and are thus prevented from being used as a template for the synthesis of complementary DNA. However, besides the gene V protein, several host functions (59, 46, 48) and the viral gene II protein (50, 51) are required in the synthesis of the viral DNA strands. Since the gene II protein is presumed to act as a strand-specific endonuclease, it is very attractive to speculate that this protein is involved in cleaving unit-length pieces of single-stranded viral DNA from the growing tail of the replicating RF molecules. In a subsequent step in phage maturation, the resulting lineair single-stranded DNA fragments must be circularized by joining the two ends of the viral strand. Although the nature of the latter mechanism is unknown, recent experiments with phage ØX-174 indicate that the cistron A protein not only introduces the nick in the viral DNA strand (60), but is also involved in ligating the single-stranded DNA fragment in order to form a covalently closed, circular molecule (61, 62). Based on these observations and on the apparent function homology between the cistron A protein of phage ØX-174 and the gene II product of the Ff phages, it is reasonable to speculate that in the case of the filamentous phages, circularization of the viral DNA strand is mediated by the gene II protein.

Finally, progeny single-stranded DNA accumulates in the infected cell as rod-like, gene V protein-coated, DNA structures which appear to be weakly associated with the inner cell membrane (63, 64). At that time also several viral proteins, such as the gene II, III, IV and VIII proteins, have accumulated at this site and here virion morphogenesis begins when the gene V protein-DNA complex starts an interaction with the accumulated amounts of the gene VIII-encoded major coat protein. The amounts of gene V and gene VIII protein involved indicate that, in the following metamorphosis, each gene V protein is replaced by two gene VIII proteins (11). Since mature virions have never been observed within infected cells (65), it is very likely that the replacement reaction occurs concomitant with the release of phages through the cell wall. Although a detailed knowledge of the latter mechanism is still lacking, it is presumed that the gene III protein, together with the gene VI protein, is involved in the termination of the phage extrusion reaction (11). It is therefore plausible to assume that the tip of the virion where the gene III protein is located is extruded last.

1.4. Genes on the Ff genome

As revealed by genetic analysis with several types of mutants, the genome of the filamentous phages contains eight complementation groups representing eight distinct genes (66, 67, 68). By recombination experiments with various amber mutants, these genes, numbered I through VIII, have been ordered in a circular array (69). Since then, the progress in the elucidation of the genetic organisation of the Ff genome was greatly accelerated by the results obtained with several types of *in vitro* studies. In this field, both Konings *et al.*, (70, 25, 71) and Model *et al.*, (72, 73) have made important contributions by using DNA-dependent cell-free protein synthesizing systems. Based upon premature termination of proteins specified by RF molecules obtained from several amber mutant phages, they have succeeded in an unambigious identification of most viral proteins. In similar studies, Model and Zinder (72) and Van den Hondel *et al.*, (74) provided definite proof that the direction of transcription on the Ff RF molecule

occurs in a counterclockwise direction on the conventional genetic map. An interesting discovery made by Konings et al., (71) was the observation that the sizes of certain in vitro synthesized proteins are not identical to the sizes of their in vivo counterparts. Both the major and the minor coat protein (i.e. the gene VIII and gene III proteins) synthesized in vivo were found to be considerably smaller than the corresponding products synthesized in vitro suggesting that these coat proteins are processed during phage maturation.

An experimental approach to the study of the Ff genome having implications which are as important as the in vitro protein synthesis studies, was the construction of several restriction enzyme cleavage maps (75, 76, 77, 78, 79, 80). With the aid of these maps it became possible to isolate specific parts of the RF molecule thus allowing detailed studies on isolated genes and on DNA regions involved in the regulation of gene expression. By using restriction fragments as DNA templates in protein-synthesizing systems (74) or in marker rescue experiments with $CaCl_2$-treated E.coli cells (81, 78) Van den Hondel et al., provided evidence that the gene order on the M13 genome is IV-(I, VI)-III-VIII-VII-V-II with II and IV being contiguous on the circular map (see figure inside front cover). Interesting features revealed by these studies are the presence of an intergenic DNA region on the Ff genome and the identification of an unknown viral polypeptide; the X protein. The noncoding intergenic region (IG) was found to be located between gene IV and gene II on the Ff genome. According to data obtained by several authors (40, 82, 83, 52, 79, 84, 78), this particular region is about 500 nucleotides long and contains the origins of replication for the synthesis of both viral and complementary DNA strands. The identification of the X-protein as a viral gene product was demonstrated by in vitro translation of restriction fragments containing the coding information for the C-terminal end of the gene II product (74). In this thesis (Chapter V), it is demonstrated that gene X is encompassed on an RNA molecule which is initiated at a promoter located within gene II. It is, therefore, most likely that gene X represents a gene within a gene. Strikingly similar observations have been made for phage ØX-174. On this

genome the DNA region encoding the C-terminal end of the cistron A protein was found to specify an additional protein, the gene B protein which has a role in phage morphogenesis (85, 86).

The various Ff-specific genes, as determined by genetic analysis as well as by *in vitro* protein synthesis studies, are listened in Table 1, together with the molecular weights and presumed functions of their encoded proteins. The molecular weights given for the proteins specified by gene I, II, III, IV and X were determined by analysing the *in vitro* synthesized products by SDS-gel electrophoresis (72, 71). As far as the gene V and gene VIII encoded proteins are concerned, their molecular weights were obtained from the estimated amino acid sequences (16, 17, 87, 88). Since the proteins specified by the genes VI and VII have to date never been detected *in vivo* or *in vitro* (73, 72, 71), the molecular weights given for these proteins have been reconstructed from recently obtained nucleotide sequence data of M13 DNA (28, 89, Van Wezenbeek *et al.*, in preparation).

TABLE 1

The genes and gene products of the Ff phages

Gene	Molecular weight of the gene product	Physiological role of the gene product
I	35,000	*morphogenesis*
II	40,000 - 46,000	*DNA replication*
III	(42,000)- 68,000	*minor capsid protein*
IV	49,000	*morphogenesis*
V	9,700	*SS synthesis*
VI	(12,350)	*morphogenesis*
VII	(3,600)	*morphogenesis*
VIII	5,200	*major capsid protein*
X	12,000	*unknown*

Values in brackets refer to data obtained by DNA sequencing studies.

1.5. Transcription of the Ff genome

Upon infection the single-stranded circular Ff genome is converted into a double-stranded, circular RF molecule. The latter, so called 'parental' RF molecule is then multiplicated yielding a pool of some 100-200 RF molecules. This pool of RF molecules serves a dual role in the life cycle of the filamentous phages; on the one hand they serve as a substrate for asymetrical viral DNA synthesis and, on the other, they are utilized as a template for the E.coli RNA polymerase in order to express the various viral genes.

In a first attempt to identify the transcription products formed on the Ff RF molecule, Takanami et al., analysed the RNA molecules transcribed from the fd RF molecule by E.coli RNA polymerase holoenzyme (90). These in vitro studies revealed that, in the transcription reaction four discrete RNA species were formed ranging in size from 10S (molecular weight $\sim 4 \times 10^5$) upto about 26S (molecular weight $\sim 1.3 \times 10^6$). Neither of these RNA species hybridised to the viral DNA strand of the RF molecule indicating that only the complementary DNA strand forms a template for RNA polymerase. Furthermore, the RNA species generated revealed different 5'-terminal nucleotide sequences. These observations led Okamoto et al., (91) to the postulation of a transcription mechanism according to which RNA synthesis on the fd RF molecule is initiated at four different promoter sites but terminated at only one, unique site.

Jacob et al., (92, 93), investigated the size distribution of phage M13-specific RNA species present in infected E.coli cells. In their approach, the total RNA population was subjected to density gradient centrifugation under denaturing conditions after which phage-specific transcripts were identified by hybridisation to denatured M13 RF. The data obtained in these studies suggest that in the infected E.coli cell at least some phage-specific RNA species are formed with sizes equivalent to the sizes of the RNA chains synthesized in vitro under the direction of replicative form DNA. Moreover, the hybridisation data obtained in both types of experiments indicated that, in vivo as well as in vitro, only the non-viral strand of the RF molecule is copied in the transcrip-

tion process which implies that under both conditions the direction of transcription along the phage genome is the same.

1.6. The aim of the present investigation

One of the most interesting features of the gene expression mechanism of Ff phages is, that *in vitro* as well as *in vivo* the proteins encoded by the genes V and VIII (representing the DNA-binding protein and the major coat protein respectively) are synthesized in much larger quantities than the other phage-encoded proteins (26, 63). Since there is probably no temporal control during expression of the Ff genome, these quantitative differences must be achieved by a constitutive regulation mechanism either on the level of transcription or on the level of translation. Based upon the *in vitro* transcription data obtained by Takanami's group, it is an attractive idea to suppose that the abundant synthesis of the gene V and gene VIII proteins is inherant in the transcription mechanism of the Ff genome. Their results suggest that the 3'-terminal regions of all *in vitro* synthesized RNA species overlap and hence, that the genes which are located proximal to (in front of) the unique termination signal for RNA synthesis are transcribed very frequently. If these terminator-proximal genes are also equipped with efficient ribosome binding sites, then these high transcription frequencies will automatically lead to high levels of expression. In a similar way one might argue that genes which are located distal to (behind) the unique transcription termination signal have low transcription frequencies and are expressed only poorly. This hypothesis, therefore, predicts that the termination signal for RNA synthesis is located immediately distal of those genes whose protein products are synthesized in the largest amounts. However, neither the *in vitro* data obtained by Takanami's group nor the *in vivo* studies by Jacob *et al.* have provided any information concerning the position of this termination signal or the codogenic capacities of the various RNA species. So, there was no evidence available to support the regulatory role of the transcription process in the expression of the Ff genome. Since there was no experimental evidence either to suppose a regulation on

translational level, the expression mechanism of the Ff genome remained completely obscure until about 1975.

The aim of the experimental work described in this thesis has been the elucidation of the mechanism(s) which regulate expression of the phage M13 genome. Although for this purpose an *in vivo* study of various phage-specific transcriptional and translational products appears to be the most direct approach, there are several reasons which militate against it. The excessive amount of host-specific material must be considered as one of the most important disadvantages of the *in vivo* system. Due to the fact that the filamentous phages do not shut off host-cell synthesis, detection and analysis of viral transcripts and proteins is seriously hampered. In order to circumvent this difficulty and to take advantage of the facilities offered by the use of restriction enzymes, we decided to use both *in vitro* transcription and translation systems as expedients to study expression of the M13 genome. Although it is always questionable to what extent *in vitro* systems reflect the *in vivo* situation, previous studies (70, 25, 71) have indicated that the cell-free protein synthesizing system used in our experiments is able to synthesize the different phage-specific proteins with great precision. As far as the RNA synthesizing system is concerned, this has been optimised for the transcription of M13 RF (Chapter II) and much of the data obtained with this system has been confirmed very recently in both DNA and RNA sequencing studies (94, 28, 89, Van Wezenbeek *et al.*, in preparation).

In the following chapter a brief survey of some aspects of the transcription phenomenon is presented. This survey is intended as an introduction to Chapters III-VIII, in which work directed towards the location of both RNA initiation and termination sites on the phage M13 genome is discussed.

In Chapter III the mapping of an efficient RNA termination signal on M13 RF is described, together with a discussion of the characterization of the RNA species formed when this template molecule undergoes transcription with *E.coli* RNA polymerase holoenzyme.

The mapping of the various RNA initiation sites on the M13

RF molecule is examined in Chapter IV. In this study use was made of several restriction enzyme cleavage maps for precise location of these sites. The relative initiating capacities of the RNA initiation sites were determined and the specific 'start' (initiation) signals of these sites were characterized.

In Chapter V it is unambiguously demonstrated that the in vitro synthesized, pppG-initiated RNA species contain partially overlapping nucleotide sequences. Based on these observations, a 'cascade-like' transcription mechanism has been postulated.

Chapter VI is devoted to a description of nucleotide sequence studies which were performed in order to elucidate the structure of the unique transcription termination signal which is operative on the M13 genome. Several aspects of the termination event in the synthesis of RNA are discussed in this chapter in connection with the nucleotide sequence data obtained.

The experiments described in Chapter VII concern the location of the A-promoters on the phage genome and the expression mechanism of the genes located immediately distal to the transcription termination signal.

The final Chapter is devoted to the effect of the termination factor rho on the transcription mechanism of M13 RF. According to the observations presented in this chapter, at least one rho-dependent termination site exists on the phage genome. The sizes of all larger RNA species are seriously restricted because of the position of this additional termination site.

REFERENCES

1. Marvin, D.A., Pigram, W.J., Wiseman, R.L., Wachtel, E.J. and Marvin, F.J. (1974). *J. Mol. Biol.*, 88, 581-600.
2. Marvin, D.A. and Hohn, B. (1969). *Bacteriol. Rev.*, 33, 172-209.
3. Meynell, G.G. and Lawn, A.M. (1968). *Nature*, 217, 1184-1186.
4. Khatoon, H., Iyer, R.V. and Iyer, V.N. (1972). *Virology*, 48, 145-155.
5. Marvin, D.A., Wiseman, R.L. and Wachtel, E.J. (1974). *J. Mol. Biol.*, 82, 121-138.
6. Minamishima, Y., Takeya, K., Ohnishi, Y and Amako, K. (1968). *J. Virol.*, 2, 208-213.
7. Kuo, T.T., Haung, T.C. and Chow, T.Y. (1969). *Virology*, 39, 548-554.
8. Loeb, T. (1960). *Science*, 131, 932-933.
9. Hofschneider, P.H. (1962). *Z. Naturforsch.*, 18b, 203-210.
10. Hoffmann-Berling, H., Marvin, D. and Dürwald, H. (1963). *Z. Naturforsch.*, 18b, 876-883.
11. Denhardt, D.T. (1975). In: *CRC Critical Reviews in Microbiology*. The Chemical Rubber Company, Cleveland, Ohio, pp. 161-223.
12. Ray, D.S. (1977). *Replication of Filamentous Bacteriophages. Comprehensive Virology*. (H. Fraenkel-Conrat and R.R. Wagner, eds.), Plenum Publishing Corporation, New York, pp. 105-178.
13. Salivar, W.O., Tzagoloff, H. and Pratt, D. (1964). *Virology*, 24, 359-371.
14. Marvin, D.A. and Schaller, H. (1966). *J. Mol. Biol.*, 15, 1-7.
15. Berkowitz, S.A. and Day, L.A. (1974). *Biochemistry*, 13, 4825-4831.
16. Asbeck, F., Beyreuther, K., Kohler, H., von Wettstein, G. and Braunitzer, G. (1969). *Hoppe-Seyler's Z. Physiol. Chem.*, 350, 1047-1066.
17. Nakashima, Y. and Koningsberg, W. (1974). *J. Mol. Biol.*, 88, 598-600.
18. Lin, J.Y., Wu, C.C. and Kuo, T.T. (1971). *Virology*, 45, 38-41.
19. Wiseman, R.L., Dunker, A.K. and Marvin, D.A. (1972). *Virology*, 48, 230-244.

20. Snell, D.T. and Offord, R.E. (1972). *Biochem. J.*, 127, 167-178.
21. Rossomando, E. (1970). *Virology*, 42, 681-687.
22. Marco, R. (1975). *Virology*, 68, 280-283.
23. Goldsmith, M.E. and Koningsberg, W.H. (1977). *Biochemistry*, 16, 2686-2694.
24. Woolford, J.L., Steinman, H.M. and Webster, R.E. (1977). *Biochemistry*, 16, 2694-2700.
25. Konings, R.N.H., Jansen, J., Cuypers, T. and Schoenmakers, J.G.G. (1973). *J. Virol.*, 12, 1466-1472.
26. Henry, T.J. and Pratt, D. (1969). *Proc. Natl. Acad. Sci. U.S.A.*, 62, 800-809.
27. Jazwinsky, S.M., Marco, R. and Kornberg, A. (1973). *Proc. Natl. Acad. Sci. U.S.A.*, 70, 205-209.
28. Schaller, H. and Takanami, M. (1978). In: *Single-Stranded DNA Phages*. (D.T. Denhardt, D.S. Ray and D. Dressler, eds.). In press.
29. Novotny, C.P. and Fives-Taylor, P. (1974). *J. Bacteriol.*, 117, 1306-1311.
30. Marco, R., Jaswinsky, S.M. and Kornberg, A. (1974). *Virology*, 62, 209-223.
31. Trenkner, E., Bonhoeffer, F. and Gierer, A. (1967). *Biochem. Biophys. Res. Commun.*, 28, 932-939.
32. Smilowitz, H. (1974). *J. Virol.*, 13, 94-99.
33. Jazwinsky, S.M., Lindberg, A.A. and Kornberg, A. (1975). *Virology*, 66, 283-293.
34. Brutlag, D., Schekman, R. and Kornberg, A. (1971). *Proc. Natl. Acad. Sci. U.S.A.*, 68, 2826-2829.
35. Fidanián, H.M. and Ray, D.S. (1972). *J. Mol. Biol.*, 72, 51-63.
36. Wickner, W., Brutlag, D., Schekman, R. and Kornberg, A. (1972). *Proc. Natl. Acad. Sci. U.S.A.*, 69, 965-969.
37. Geider, K. and Kornberg, A. (1974). *J. Biol. Chem.*, 249, 3999-4005.
38. Schekman, R., Weiner, A. and Kornberg, A. (1974). *Science*, 186, 987-993.
39. Geider, K., Beck, E. and Schaller, H. (1978). *Proc. Natl. Acad. Sci. U.S.A.*, 75, 645-649.

40. Tabak, H.F., Griffith, J., Geider, K., Schaller H. and Kornberg, A. (1974). *J. Biol. Chem.*, 249, 3049-3054.
41. Marians, J.K., Ikeda, J., Schlagman, S. and Hurwitz, J. (1977). *Proc. Natl. Acad. Sci. U.S.A.*, 74, 1965-1968.
42. Yasumoto, C.S. and Hurwitz, J. (1977). *Proc. Natl. Acad. Sci. U.S.A.*, 74, 4195-4199.
43. Hohn, B., Lechner, H. and Marvin, D.A. (1971). *J. Mol. Biol.*, 56, 143-154.
44. Olsen, W.L., Staudenbauer, W.L. and Hofschneider, P.H. (1972). *Proc. Natl. Acad. Sci. U.S.A.*, 69, 2570-2573.
45. Staudenbauer, W.L., Olsen, W.L. and Hofschneider, P.H. (1973). *Eur. J. Biochem.*, 32, 247-253.
46. Staudenbauer, W.L. (1974). *Eur. J. Biochem.*, 42, 249-256.
47. Fidanián, H.M. and Ray, D.S. (1974). *J. Mol. Biol.*, 83, 63-82.
48. Ray, D.S., Dueber, J. and Suggs, S. (1975). *J. Virol.*, 16, 348-355.
49. Pratt, D. and Erdahl, W.S. (1968). *J. Mol. Biol.*, 37, 181-200.
50. Lin, N. and Pratt, D. (1972). *J. Mol. Biol.*, 72, 37-49.
51. Tseng, B.Y. and Marvin, D.A. (1972). *Virology*, 10, 384-391.
52. Horiuchi, K. and Zinder, N.D. (1976). *Proc. Natl. Acad. Sci. U.S.A.*, 73, 2341-2345.
53. Van den Hondel, C.A. (1976). *Ph.D. Thesis*, University of Nijmegen.
54. Mazur, B.J. and Model, P. (1973). *J. Mol. Biol.*, 78, 285-300.
55. Mazur, B.J. and Zinder, N.D. (1975). *Virology*, 68, 490-502.
56. Alberts, B., Frey, L. and Delius, H. (1972). *J. Mol. Biol.*, 68, 139-152.
57. Salstrom, J.S. and Pratt, D. (1971). *J. Mol. Biol.*, 61, 489-501.
58. Oey, J.L. and Knippers, R. (1972). *J. Mol. Biol.*, 68, 125-136.
59. Mitra, S. and Stallions, D.R. (1973). *Virology*, 52, 417-424.
60. Langeveld, S.A., van Mansfield, A.D.M., Baas, P.D., Jansz, H.S., van Arkel, G.A. and Weisbeek, P.H. (1968). *Nature*, 271, 417-420.
61. Eisenberg, S., Griffith, J. and Kornberg, A. (1977). *Proc. Natl. Acad. Sci. U.S.A.*, 74, 3198-3202.

62. Baas, P.D. and Jansz, H.S. (1978). In: *Single-Stranded DNA Phages*. (D.T. Denhardt, D.S. Ray and D. Dressler, eds.) Cold Spring Harbor Laboratory, in press.
63. Webster, R.E. and Cashman, J.S. (1973). *Virology*, 55, 20-38.
64. Pratt, D., Laws, P. and Griffith, J. (1974). *J. Mol. Biol.*, 82, 425-439.
65. Stegen, U. and Hofschneider, P.H. (1970). *J. Mol. Biol.*, 48, 361-364.
66. Pratt, D., Tzagoloff, H. and Erdahl, W.S. (1966). *Virology*, 30, 397-410.
67. Mazur, B.J. and Zinder, N.D. (1975). *Virology*, 68, 284-285.
68. Pratt, D., Tzagoloff, H. and Beaudoin, J. (1969). *Virology*, 39, 42-53.
69. Lyons, L.B. and Zinder, N.D. (1972). *Virology*, 49, 45-60.
70. Konings, R.N.H. (1973). *FEBS Lett.*, 35, 155-160.
71. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975). *J. Virol.*, 15, 570-584.
72. Model, P. and Zinder, N.D. (1974). *J. Mol. Biol.*, 83, 231-251.
73. Model, P., Horiuchi, K., McGill, C. and Zinder, N.D. (1975). *Nature (London)*, 253, 132-134.
74. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Virology*, 67, 487-497.
75. Takanami, M. (1973). *FEBS Lett.*, 34, 318-322.
76. Takanami, M., Okamoto, T., Sugimoto, K. and Sugisaki, H. (1975). *J. Mol. Biol.*, 95, 21-31.
77. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 547-558.
78. Van den Hondel, C.A., Pennings, L. and Schoenmakers, J.G.G. (1976). *Eur. J. Biochem.*, 68, 55-70.
79. Vovis, G.F., Horiuchi, K. and Zinder, N.D. (1975). *J. Virol.*, 16, 674-684.
80. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1976). *J. Virol.*, 18, 1024-1039.
81. Van den Hondel, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 559-567.
82. Suggs, S. and Ray, D.S. (1977). *J. Mol. Biol.*, 110, 147-164.

83. Gray, C.P., Sommer, R., Polke, C., Beck, E. and Schaller, H. (1978). *Proc. Natl. Acad. Sci. U.S.A.*, 75, 50-53.
84. Schaller, H., Uhlman, A. and Geider, K. (1976). *Proc. Natl. Acad. Sci. U.S.A.*, 73, 49-53.
85. Weisbeek, P.J., Borrias, W.E., Langeveld, S.A., Baas, P.D. and van Arkel, G.A. (1977). *Proc. Natl. Acad. Sci. U.S.A.*, 74, 2504-2508.
86. Sanger, F., Air, G.M., Barrell, B.G., Brown, A.L., Coulson, A.R., Fiddes, J.C., Hutchinson, C.A., Slocombe, P.M. and Smith, M., *Nature*, 265, 687-695.
87. Nakashima, Y., Dunker, A.K., Marvin, D.A. and Koningsberg, W. (1974). *FEBS Lett.*, 43, 125.
88. Cuypers, T., van der Ouderaa, F.J. and de Jong, W.W. (1974). *Biochem. Biophys. Res. Commun.*, 5962, 557-563.
89. Hulsebos, T. and Schoenmakers, J.G.G. (1978). Submitted.
90. Takanami, M., Okamoto, T. and Sugiura, M. (1970). *Cold Spring Harbor Symp. Quant. Biol.*, 34, 179-187.
91. Okamoto, T., Sugiura, M. and Takanami, M. (1969). *J. Mol. Biol.*, 45, 101-111.
92. Jacob, E., Jaenisch, R. and Hofschneider, P.H. (1970). *Nature (London)*, 227, 59-60.
93. Jacob, E., Jaenisch, R. and Hofschneider, P.H. (1973). *Eur. J. Biochem.*, 32, 432-443.
94. Rivera, M.J., Smits, M.A., Quint, W., Schoenmakers, J.G.G. and Konings, R.N.H. (1978). *Nucleic Acad Res.*, 5, 2895-2912.

CHAPTER II

SOME ASPECTS OF THE TRANSCRIPTION PHENOMENON

2.1. An outline of the transcription process

Transcription in bacterial cells is mediated by DNA-dependent RNA polymerase. This enzyme has a molecular weight of about 500,000 and contains four major subunits. These subunits are designated β, β', σ and α and are present in the molar ratio 1 : 1 : 1 : 2 (1, 2). In the presence of a suitable DNA template, the four ribonucleoside triphosphates, Mg^{2+}, and under appropriate ionic conditions, RNA polymerase holoenzyme (i.e. RNA polymerase containing all four subunits) is capable of synthesizing discrete RNA products. Since the synthesis of discrete products implies proper initiation and termination events, both the RNA initiation and the RNA termination event must be signaled by a direct interaction between the template DNA and the RNA polymerase. According to current models (3, 4) the various steps in the transcription process involve:

1. recognition and template binding of the RNA polymerase. The enzyme forms a complex with the template DNA at a specific recognition sequence. This recognition sequence forms part of the promoter and is located upstream of the RNA-initiation site. Although it is generally assumed that in this stage binding of the enzyme occurs at the outside of the helix, recent evidence challenges this view and suggests that opening of base pairs is also essential for promoter selection (5).

2. formation of a pre-initiation complex. The transition from the recognition complex to the stable pre-initiation complex involves melting of the DNA strands over about 10-15 base pairs (6) and is therefore strongly dependent on salt concentration and temperature. Based on RNA polymerase binding sequences of several DNA genomes, the sequence 5' TATPuATG 3'
3' ATAPyTAC 5' (also referred to as 'Pribnow' box) is involved in this process (4). Initiation of RNA synthesis occurs usually about 5 or 6 nucleotides beyond this sequence.

3. initiation. With the RNA polymerase in the partially unwound DNA helix, the first ribonucleoside triphosphate (usually ATP or GTP) is introduced in the enzyme - DNA complex. After the formation of a phosphodiester bond between the first and the second template-encoded nucleotide, a dinucleoside tetraphosphate of the structure pppPupX is formed and initiation of RNA synthesis is effectively complete.

4. RNA chain elongation. Successive ribonucleoside monophosphate residues are added from the substrate ribonucleoside triphosphates to the 3'-OH end of the nascent RNA chain and inorganic pyrophosphate is released. The RNA polymerase specificity subunit sigma, which has played an important role in the recognition step, is released from the complex (7).

5. RNA chain termination and enzyme release. As a result of the transcription of a 'termination' sequence in the template DNA, RNA synthesis ceases and both the completed RNA chain and the RNA polymerase 'core' (i.e. RNA polymerase without the sigma subunit) are released from the template. Features of the transcription termination event will be discussed in Chapter VI on the basis of the nucleotide sequence of a termination signal for RNA synthesis on the phage M13 genome.

2.2. Regulatory factors in the transcription process

Although initiation and termination of transcription is brought about by a direct interaction between RNA polymerase holoenzyme and the DNA template, the efficiency with which these two events occur is sometimes regulated by additional factors. Classic and well-known examples of templates, the transcription of which is regulated on the level of RNA initiation, are the immunity region on the genome of phage lambda (8, 9, 10) and the lactose (*lac*) operon of E.coli (11, 12, 13, 14). In case of the (*lac*) operon the frequency of RNA initiation is regulated by two mechanisms; the *lac* repressor system which represents a negative control and the catabolite gene activator protein (CAP) representing a positive control.

The *lac* repressor is able to switch off transcription of the structural genes by binding to the operator and thus preventing initiation of transcription at the *lac* promoter. On the other hand, the CAP factor is able to stimulate initiation of transcription since CAP factor activated by cyclic AMP and bound to the CAP interaction site facilitates recognition and binding of RNA polymerase to the *lac* promoter. Whether regulatory sites on the level of RNA initiation are also operative in combination with promoters which are located on the genomes of the Ff phages is unclear.

Evidence from in vitro and in vivo studies suggests that besides the RNA initiation event the RNA termination event is in many instances regulated by additional factors. Of such factors the rho protein, which is encoded by the Su-A gene of E.coli (15,16) has been the most extensively studied (for a review see 17). Although the mechanism by which rho interferes with the transcription process is still obscure, it is known that the protein catalyses termination of transcription at specific, rho-dependent termination sites. At these sites, which have been shown to occur at various templates such as the RF molecules of the filamentous phages (18, Chapter VIII), RNA synthesis ceases when rho is present. In the absence of rho, termination of transcription at rho-dependent termination sites does not occur or occurs with a very low efficiency (19, 20).

The best characterized effects of rho on RNA synthesis in vitro in relation to the transcription process in vivo have been found to occur on the DNA template of phage lambda (21, 22, 17 23) and on the tryptophan (*trp*) operon of E.coli (24,17). On both templates rho is known to stimulate termination of transcription at sites which are also involved in vivo in the termination process. In the *trp* operon of E.coli rho influences the transcription frequency of the structural genes by regulating the efficiency of a transcription termination site (the 'attenuator') which is located immediately proximal to the first structural gene. On the genome of bacteriophage lambda, rho is involved in termination events which regulate the length of transcripts initiated within the immunity region at the promoters P_r and P_l. The latter two promoters (controlled by the operators O_r and O_l)

largely govern the initiation of the lytic pathway of the phage.

2.3. Conditions affecting the *in vitro* transcription process

In studying the interaction between RNA polymerase and DNA templates, *in vitro* transcription systems have proved to be valuable tools. However, the *in vitro* reaction is influenced by several factors which interfere with the selectivity of the transcription process (for a review see 3). The most striking alterations are due to changes in the RNA polymerase or in the DNA template. Due to its role in promoter selection, removal or inactivation of the sigma subunit of the RNA polymerase has a profound effect on the selectivity of the transcription reaction. It has been reported that RNA polymerase preparations without sigma factor are unable to locate promoter sites or to open the DNA strands in order to initiate transcription (25, 26, 27).In addition, the strand specificity of the transcription reaction is lost en the enzyme initiates randomly along the template (28, 27, 12). Random initiation has also been reported when the template DNA is completely denatured (29,30,31).In this case the DNA strands are transcribed completely forming small RNA fragments. When single-strand nicks are introduced in double-stranded helical DNA templates (the M13 RF molecule for example), the selectivity of the transcription reaction is barely diminished, though the amount of RNA formed in the transcription reaction is seriously reduced (32, 5).Besides the quality of the RNA polymerase and the DNA template, factors such as the concentration of the ribonucleoside triphosphates and the ionic strength of the reaction mixture influence the *in vitro* transcription process. As has been observed by Minkley and Pribnow (33), very little initiation takes place when all RNA precursors are present at concentrations of 5 μM or lower. In addition, such low concentrations of triphosphate can cause pausing of the RNA polymerase during chain elongation (34). Although the synthesis of RNA from various templates and with RNA polymerases from various bacteria shows great differences in sensitivity to ionic strength, elevated salt concentrations have been reported to alter the process of promoter selection (35, 5), the rate of RNA chain initiation

(35) the rate of chain elongation (36, 37, 38) and the efficiency of the termination event (37, 39, 36). As far as the filamentous phages are concerned, our (unpublished) results have shown that varying the KCl concentration from 50 mM up to 200 mM hardly affects the sizes, the number and the relative amounts of the various phage-specific RNA species. These observations, therefore, lend support to the notion that the promoters on the phage M13 genome do not exhibit large differences in their sensitivity to a variation in ionic strength. However, KCl concentrations lower than 50 mM and higher than 150 mM reduce the total amount of RNA synthesized, indicating that the ionic strength does influence transcription of this template molecule.

Although the reason for the effect of KCl on the transcription reaction is not well known, it is thought that KCl stabilizes the DNA. This stabilisation results in a reduced efficiency of strand separation and thus interferes with the interaction between RNA polymerase and the template DNA. The end effect on the level of RNA initiation is that fewer initiation complexes are formed on the DNA and hence that the specificity of binding is increased. The effect of the KCl concentration in enhancing the efficiency of the termination event may be explained by assuming that high KCl concentrations favor spontaneous dissociation of the enzyme-DNA-RNA complex and therefore facilitate termination of transcription. The latter effect is, however, in contrast with the effect of KCl on rho-mediated transcription termination events. In this case low KCl concentrations (< 100mM) stimulate the termination event presumably, because rho functions optimally at low ionic strength (22, 18).

2.4. Techniques for the localization of promoters

Due to their specific capacities, promoter sites can be mapped on DNA genomes by several *in vitro* techniques. These techniques include:

1. mapping of restriction fragments which are able to form a firm complex with RNA polymerase (40, 41, 42, 43, 44).

2. mapping of restriction fragments which are able to initiate RNA synthesis (42, Chapter IV).

3. mapping of restriction fragments which are radioactively labeled upon hybridisation to 5'-terminally-labeled RNA (45, 46).

4. localization of DNA regions able to bind RNA polymerase using electron microscopy (47, 48, 49, 50).

5. identification of polypeptides encoded by isolated RNA species in an RNA-dependent cell-free protein synthesizing system (51, Chapters III and V).

6. identification of restriction fragments which are able to direct the synthesis of polypeptides in a DNA-dependent cell-free protein synthesizing system (52, Chapters V and VII).

A technique which has had important implications for the localisation of promoters as well as for nucleotide sequence analysis of promoters, has been developed by Schaller and coworkers. In this technique (mentioned under one above),RNA polymerase holoenzyme and DNA restriction fragments are allowed to form stable complexes in the absence of ribonucleoside triphosphates. Since DNA-RNA polymerase complexes are quantitatively retained by membrane filters, whereas RNA polymerase and native DNA separately pass through the filter, restriction fragments containing promoters can be identified. In order to obtain information concerning the nucleotide sequences which are able to bind RNA polymerase, a slight modification of this technique has been used. After binding of the RNA polymerase to the DNA fragment, the enzyme-DNA complex is subjected to digestion with pancreatic DNase. Since the RNA polymerase binding sequence is protected from digestion, the small DNA fragment containing this sequence can be isolated (40, 25, 4, 53, 54, 55, 56). The latter fragments are usually about 40 base pairs long and contain the characteristic sequence (Pribnow box) which is involved in the RNA polymerase binding event.

In our approach to the location of promoters on the phage M13 genome, we have used the technique listened under two above. In this technique, isolated restriction fragments of the RF molecule have been used as templates in an *in vitro* RNA synthesizing system. Since, under appropriate transcription conditions, only those fragments which contain promoters are able to initiate RNA synthesis, promoter sites can be located on particular DNA fragments. By estimating the length of the RNA product synthesized, it is also possible to map the promoter on the restriction fragment. From the size of the DNA fragment, the direction of transcription and the length of the RNA chain, the promoter position on the fragment can be calculated exactly. One assumption made in these calculations is that termination of transcription occurs at the end of the restriction fragment. Although this assumption has shown to be correct, we have noted that in some cases RNA polymerase holoenzyme tends to proceed the polymerisation reaction by switching to the complementary DNA strand of the restriction fragment (Chapters IV and V). This aberrant behaviour of the enzyme is especially stimulated under low salt conditions (< 100 mM KCl; M. Takanami, personal communications; our unpublished results), thus providing direct evidence that a low ionic strength the selectivity of transcription is reduced.

Other techniques we have used to locate promoter sites are listened under five and six. In these approaches either isolated RNA species or isolated restriction fragments have been used to programme cell-free protein synthesizing-systems. Whereas the transcription studies mentioned above provided information concerning promoter locations on restriction enzyme cleavage maps, the two types of translation studies enable the location of promoters on the genetic map of the phage genome. When isolated RNA species are used to direct the protein synthesizing system, promoter locations can be deduced from the coding information of these transcripts since the order of genes and the direction of transcription on the phage genome are known. When purified restriction fragments are used in the protein synthesising system, the expression of genes encompassed on the fragment is dependent upon the existance of promoters on these fragments. If, therefore, a restriction fragment is able to direct the synthesis of a parti-

cular protein, the fragment must contain a promoter preceding the gene coding for that protein.

REFERENCES

1. Burgess, R. (1969). *J. Biol. Chem.*, 244, 6168-6176.
2. Burgess, R., Travers, A., Dunn, J., Bautz, E. (1969). *Nature*, 221, 43-46.
3. Chamberlin, M.J. (1974). *Annu. Rev. Biochem.*, 43, 721-775.
4. Pribnow, D. (1975). *J. Mol. Biol.*, 99, 419-443.
5. Seeburg, P.H., Nüsslein, C. and Schaller, H. (1977). *Eur. J. Biochem.*, 74, 107-113.
6. Saucier, J. and Wang, J. (1972). *Nature New Biol.*, 239, 167-170.
7. Wu, C.W., Yarbrough, L.R. and Hillel, Z. (1975). *Proc. Natl. Acad. Sci. U.S.A.*, 72, 3019-3023.
8. Maniatis, T., Ptashne, M., Backman, K., Kleid, D., Flashman, S., Jeffrey, A. and Maurer, R. (1975). *Cell*, 5, 109-113.
9. Ptashne, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B. and Sauer, R.T. (1976). *Science*, 194, 156-161.
10. Waltz, A., Pirrotta, V. and Ineichen, K. (1976). *Nature*, 262, 665-669.
11. Dickson, R.C., Abelson, J., Barnes, W.M., Reznikoff, W.S. (1975). *Science*, 187, 27-35.
12. Eron, L. and Block, R. (1971). *Proc. Natl. Acad. Sci. U.S.A.*, 68, 1828-1832.
13. Majors, J. (1975). *Nature*, 256, 672-674.
14. Majors, J. (1975). *Proc. Natl. Acad. Sci. U.S.A.*, 72, 4394-4398.
15. Bachman, J., Low, K. and Taylor, A. (1976). *Bact. Rev.*, 40, 116-167.
16. Ratner, D. (1976). In: *RNA Polymerase*. (R. Losick and M. Chamberlin, eds.) Cold Spring Harbor Laboratory, pp. 645-655.
17. Roberts, J.W. (1976). In: *RNA Polymerase*. R. Losick and M. Chamberlin, eds.) Cold Spring Harbor Laboratory, pp. 247-271.
18. Takanami, M., Okamoto, T. and Sugiura, M. (1971). *J. Mol. Biol.*, 62, 81-94.

19. Howard, B.H., de Crombrugghe, B. and Rosenberg, M. (1977). *Nucleic Acid Res.*, 4, 827-842.
20. Kupper, H., Sekiya, T., Rosenberg, M., Egan, J. and Landy, A. (1978). *Nature*, 272, 423-428.
21. Roberts, J. (1969). *Nature*, 224, 1168-1174.
22. Roberts, J. (1970). *Cold Spring Harbor Symp. Quant. Biol.*, 35, 121-126.
23. Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978). *Nature*, 272, 414-423.
24. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C.L., Squires, C. and Yanofsky, C. (1975). *Science*, 189, 22-26.
25. Le Talaer, J.Y. and Jeanteur, P.L. (1971). *Proc. Natl. Acad. Sci. U.S.A.*, 68, 3211-3215.
26. Hinkle, D. and Chamberlin, M. (1970). *Cold Spring Harbor Symp. Quant. Biol.*, 35, 65-72.
27. Hinkle, D., Ring, J. and Chamberlin, M. (1972). *J. Mol. Biol.*, 70, 197-207.
28. Sugiura, M., Okamoto, T. and Takanami, M. (1970). *Nature*, 225, 598-600.
29. Chamberlin, M. and Berg, P. (1964). *J. Mol.Biol.*, 8, 297-313.
30. Chamberlin, M. (1970). *Cold Spring Harbor Symp. Quant. Biol.*, 35, 851-873.
31. Maitra, U., Nakata, Y., Hurwitz, J. (1967). *J. Biol. Chem.*, 242, 4908-4913.
32. Hayashi, Y. and Hayashi, M. (1971). *Biochemistry*, 10, 4212-4218.
33. Minkley, E.G. and Pribnow, D. (1973). *J. Mol. Biol.*, 77, 255-277.
34. Maizels, N.M. (1973). *Proc. Natl. Acad. Sci. U.S.A.*, 70, 3583-3589.
35. Mangel, W. and Chamberlin, M.J. (1974). *J. Biol. Chem.*, 249, 3002-3006.
36. Fuchs, E., Milette, R.L., Zillig, W. and Walter, G. (1967). *Eur. J. Biochem.*, 3, 183-193.
37. Richardson, J.P. (1970). *Nature*, 225, 1109-1112.
38. Rhodes, G. and Chamberlin, M.J. (1974). *J. Biol. Chem.*, 349, 6675-6683.

39. Schäfer, R. and Zillig, W. (1973). *Eur. J. Biochem.*, 33, 215-226.
40. Chen, C.Y., Hutchinson, C.A. and Edgell, M.H. (1973). *Nature New Biology*, 243, 233-236.
41. Humphries, P., McConnel, D.J., Conolly, P. (1974). *Biochem. Biophys. Acta*, 366, 70-78.
42. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, M. (1975). *J. Mol. Biol.*, 95, 33-34.
43. Seeburg, P. and Schaller, H. (1975). *J. Mol. Biol.*, 92, 261-277.
44. Takeya, T. and Takanami, M. (1974). *Biochemistry*, 13, 5388-5394.
45. Axelrod, N. (1976). *J. Mol. Biol.*, 108, 771-779.
46. Smith, L.H. and Sinnsheimer, R.L. (1976). *J. Mol. Biol.*, 103, 699-710.
47. Giaccomoni, P.U., Delain, E. and Le Pecq, J.B. (1977). *Eur. J. Biochem.*, 78, 205-213.
48. Bordier, C. and Dubochet, J. (1974). *Eur. J. Biochem.*, 44, 617-624.
49. Portmann, R., Sogo, J.M., Koller, T.L. and Zillig, W. (1974). *FEBS Lett.*, 45, 64-67.
50. Koller, Th., Sogo, J.M. and Bujard, H. (1974). *Biopolymers*, 13, 995-1009.
51. Chan, T., Model, P. and Zinder, N.D. (1975). *J. Mol. Biol.*, 99, 369-382.
52. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Virology*, 67, 487-497.
53. Schaller, H., Gray, C. and Herrmann, K. (1975). *Proc. Natl. Acad. Sci. U.S.A.*, 72, 737-741.
54. Heyden, B., Nüsslein, C. and Schaller, H. (1972). *Nature New Biology*, 240, 9-12.
55. Heyden, B., Nüsslein, C. and Schaller, H. (1975). *Eur. J. Biochem.*, 55, 147-155.
56. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1975). *Nature*, 253, 410-414.

CHAPTER III

PHYSICAL MAPPING OF THE CENTRAL TERMINATOR FOR TRANSCRIPTION ON THE BACTERIOPHAGE M13 GENOME

PHYSICAL MAPPING OF THE CENTRAL TERMINATOR FOR TRANSCRIPTION ON THE BACTERIOPHAGE M13 GENOME

ABSTRACT

With the aid of *in vitro* transcription and translation studies it has been demonstrated that termination of transcription on bacteriophage M13 replicative form DNA occurs at a unique site which is located immediately distal to the 3'-end of gene VIII, the gene which codes for the major capsid protein. The position of this site has been mapped accurately on the enzyme cleavage maps by transcription of restriction fragments of M13 RF DNA. The central termination site was found to be located in restriction fragment Hap-B_2 at 450 nucleotides from the 5'-end of its viral strand (0.77 fractional length clockwise from the unique *Hind* II enzyme cleavage site).

INTRODUCTION

Transcription *in vitro* of the circular double-stranded replicative form DNA of the small filamentous bacteriophages (such as M13, fd and f1) with RNA polymerase from *Escherichia coli* results in the formation of a discrete number of RNA species ranging in size from about 8S upto 26S[1,2]. All RNA species contain a unique nucleotide sequence at their 5'-end[3,4]. Since the longest RNA, viz. 26S RNA represents nearly one round of transcription of the DNA template and the RNA polymerase has been shown to copy only the 'non-viral' strand of RF DNA, it has been postulated that the different RNA classes are the result of a transcription proces according to which RNA synthesis is initiated at different promoter sites but is terminated at the same unique termination signal[1,5].

Such a model implies, however, that all RNA species contain identical nucleotide sequences at their 3'-terminal end. Moreover, all phage M13 mRNA species should share the coding information for a protein the gene of which is located most proximal to the central termination site. Recent *in vitro* protein synthesis studies, which are directed by the various M13 RNA species synthesized *in vitro*[2] as wel as *in vivo*[21], have demonstrated that all size classes of RNA are able to synthesize the protein encoded by gene VIII. From these results it is inferred that the

central termination site is located immediately distal to this gene.

As a part of our studies on the regulation of the expression of the M13 genome we are seeking to determine the DNA nucleotide sequence which signals for this transcriptional termination. Recently, the genetic map of M13 has been related to several restriction enzyme cleavage maps [7-10]. Consequently, DNA restriction fragments which are presumed to contain this central terminator are now available[14]. For this reason we have studied the transcriptional properties of these fragments in order to map, more accurately, this termination signal. Our results suggest that termination of transcription on M13 RF DNA occurs at a site which is located approximately 450 nucleotides from the left-hand side of restriction fragment Hap-B_2 (figs. 2 and 5).

MATERIALS AND METHODS

Phage M13 RF DNA was prepared as described previously[19]. *E.coli* RNA polymerase holoenzyme was a generous gift of Dr. R. Schilperoort (University of Leiden). Methods for the preparation and purification of restriction fragments of M13 RF DNA have been described previously[7,8].

RNA synthesis *in vitro* was performed in a standard reaction mixture (0.05 ml) which contained : Tris-HCl (pH 7.9), 2µmol; KCl, 7.5 µmol; $MgCl_2$, 0.4 µmol; dithiothreitol, 0.05 µmol; EDTA, 0.05 µmol; Tween-80 0.1%; bovine serum albumin, 12.5 µg; M13 RF-I DNA, 0.1 pmol and *E.coli* RNA polymerase, 1.0 pmol. After a preincubation period of 5 min at 37°C, ribonucleoside triphosphates, one of which was (α-^{32}P) labeled UTP, were added to a final concentration of 80 µM. Initiation of RNA synthesis was allowed to proceed for one minute after which time rifampicin was added (final concentration 25µg/ml).

In experiments using either (γ-^{32}P) labeled ATP or GTP, initiation was allowed to proceed for 3 min at 37°C. In these particular cases the concentration of the triphosphates was 6 µM for the labeled and 80 µM for the non-labeled ones. Prior to the addition of rifampicin the concentration of GTP and ATP was adjusted to 80 µM. After the addition of rifampicin, RNA

synthesis was continued for 5 min at 37°C.

Restriction fragments were transcribed under similar reaction conditions with the exception that the molar ratio of RNA polymerase to DNA template was 40.

After incubation the reaction mixtures were extracted with phenol and the RNA was precipitated twice with 2.5 volumes of ethanol at -20°C. Finally, the RNA was collected by centrifugation (15 min at 150.000 g) and dissolved in 20 µl of pure formamide. After dissolution the RNA was analysed by electrophoresis for 4 h at 30 mA (about 150 V) on vertical slab gels (16 cm x 14 cm x 0.2 cm), containing 3.8% polyacrylamide in 98% formamide[12]. After electrophoresis the wet gel slab was exposed to X-ray film (Kodak, RP/R54).

For *in vitro* protein synthesis studies, the radioactive bands were cut from the gel and the RNA was extracted electrophoretically[8]. Occasionally, the RNA was further purified by CF11-cellulose chromotography as described by Franklin[22]. Cell-free protein synthesis, under the direction of the *in vitro* synthesized phage M13 mRNA species, was carried out according to the method of Konings *et. al.*[11]

RESULTS AND DISCUSSION

Transcription *in vitro* of bacteriophage M13 replicative form DNA by *E.coli* RNA polymerase holoenzyme results in the formation of at least seven RNA species, ranging in size from about 8S upto 26S. These can be separated readily by polyacrylamide gel electrophoresis (Fig. 1b). Experiments performed in the presence of (γ-^{32}P) labeled ribonucleoside triphosphates revealed that two of these RNA species (23S and 26S) are initiated with pppA (Fig. 1C; A-start) while the other five (8S, 11S, 14S, 17S and 19S) are initiated preferentially with pppG (Fig. 1d; G-start).

ABBREVIATIONS used: RF, refers to covalently closed double-stranded replicative form DNA. *Hind*, *Hae*, *Hap* and *Alu* refer to fragments obtained after cleavage of M13 RF DNA with the restriction endonucleases from *Haemophilus influenzae* (endo R. Hind II), *H.aegyptius* (endo R. Hae III), *H.aphirophilus* (endo R. Hap II) and *A.luteus* (endo R. Alu), respectively.

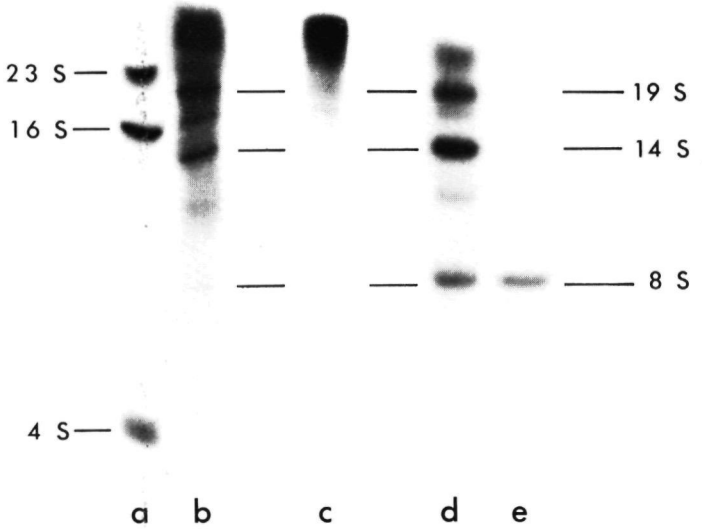


Fig. 1 : Autoradiogram of the RNA species synthesized *in vitro* under the direction of either M13 RF DNA or restriction fragment Hap-B_2. The RNA products were analysed on 3.8% polyacrylamide slab gels containing 98% formamide (12).
(a) : *E.coli* ribosomal RNA markers.
(b), (c), and (d) : The RNA products synthesized under the direction of M13 RF DNA in the presence of (α-^{32}P)-UTP, (γ-^{32}P)-ATP and (γ-^{32}P)-GTP.
(e) : The RNA products synthesized under the direction of fragment Hap-B_2 in the presence of (γ-^{32}P)-GTP.

Furthermore, the results clearly demonstrate that the RNA species with a size of 8S, 14S and 19S are synthesized in much larger quantities than the other RNA transcripts (cf. Fig. 1d).

The synthesis of a discrete number of well-defined RNA species implies that an equivalent number of promoter sites (RNA initiation sites) exist on the M13 genome. Recently, three of these promoter sites have been localized on the genetic map with the aid of coupled *in vitro* transcription and translation studies[13,14]. Furthermore, on the basis of the size of RNA chains synthesized on restriction fragments, we have been able to localize, more precisely, these and the other phage promoters on the restriction enzyme cleavage maps[2]. The positional pattern of M13 phage promoters detected, together with their respective

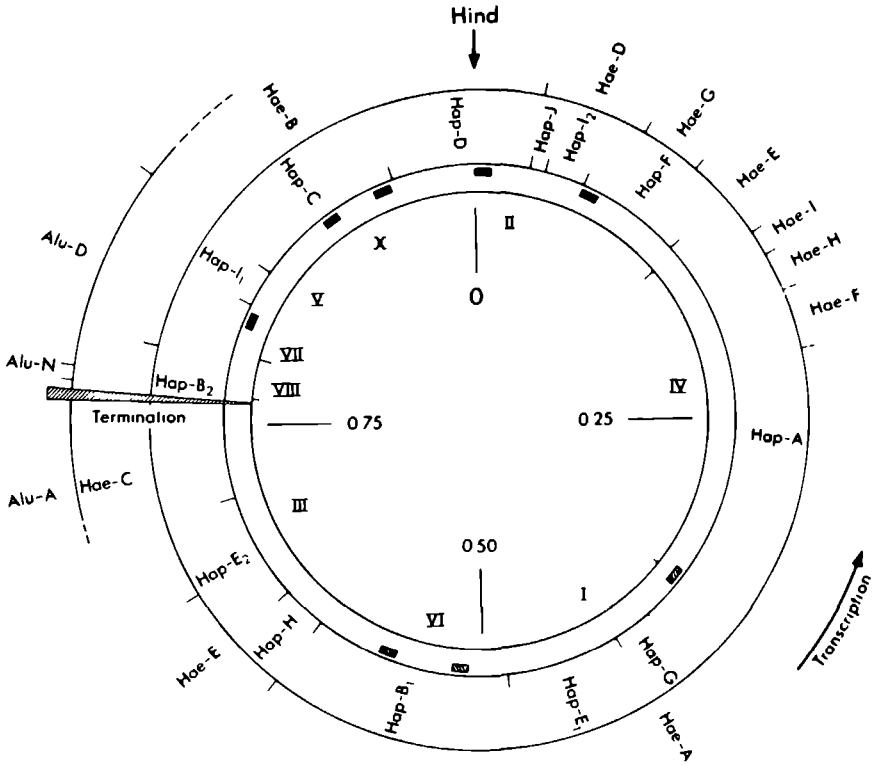


Fig. 2 : Genetic map and physical maps of bacteriophage M13. The inner circle represents the genetic map (8). The middle and outer circles show the respective locations of the Hap II- and Hae III fragments (7). Part of the physical map, obtained after cleavage of RF DNA with the Alu enzyme, is indicated by the broken circle. The positions of the 'G-promoter' are indicated with black bars, the positions of the 'A-promoter' are indicated with hatched bars. The approximate position of the central terminator for transcription is indicated.

initiation signals, are presented in Fig. 2. A generally similar distribution of RNA polymerase binding sites along the physical maps has been reported recently by Seeburg and Schaller[10].

As outlined in the Introduction in vitro protein synthesis studies directed by phage M13 specific mRNA's, which are synthesized both in vitro[2] as well as in vivo[21], have indicated that the termination site for transcription is located immediately after gene VIII, the gene which codes for the major capsid pro-

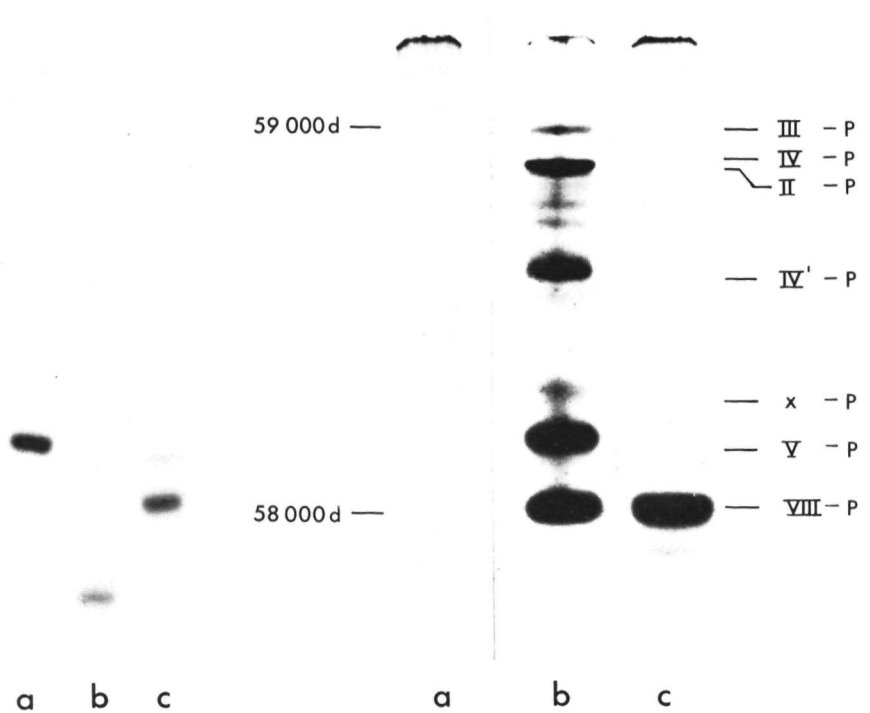


Fig. 3

Fig. 4

Fig. 3 : Autoradiogram of the RNA species synthesized in vitro under the direction of (a) : fragment Hap-B_2; (b) : fragment Hap-B_2·Hae-B; (c) : fragment Alu-D. RNA synthesis was carried out in the presence of (α-^{32}P)UTP[2]. The arrow refers to the position of migration of E.coli 16S ribosomal RNA.

Fig. 4 : Autoradiogram of the polypeptides synthesized in vitro under the direction of the in vitro synthesized phage M13 RNA's and separated by SDS-polyacrylamide gel electrophoresis. (a) : products synthesized in the absence of added RNA; (b) : products synthesized in the presence of the mRNA's transcribed from M13 RF DNA; (c) : products synthesized in the presence of the mRNA (8S RNA) transcribed from fragment Hap-B_2.

tein[6]. In order to substantiate this observation, we have studied in more detail the RNA synthesis under the direction of restriction fragment Hap-B_2 (800 base pairs[7]; Fig. 2). Previously, this fragment has been shown to direct, in a DNA-coupled in vitro protein synthesizing system, the synthesis of gene VIII protein[14]. As is demonstrated in Figure 1e, fragment Hap-B_2 directs the in

vitro synthesis of an RNA species which is approximately 360 nucleotides (8S) long. The size of this RNA is exactly the size of the smallest RNA transcribed from the intact M13 genome (Fig. 1d). In addition, both RNA chains are preferentially initiated with pppG. Furthermore, as one should expect on the basis of the results published elsewhere[2,14,21], both the smallest transcriptional product derived from the intact M13 genome (data not shown) and the 8S RNA species transcribed from fragment Hap-B_2 direct in vitro only the synthesis of the protein encoded by gene VIII (Fig. 4c). Hence it is reasonably certain that both RNA transcripts are initiated and terminated at the same functional sites of the M13 genome.

Cleavage of restriction fragment Hap-B_2 with the Hae III restriction endonuclease resulted in the formation of two new fragments[7], designated Hap-B_2·Hae-B and Hap-B_2·Hae-C, with respective lengths of approximately 300 and 500 base pairs (Figs. 2 and 5). The former fragment, which is derived from the left-hand side ('5'-end') of fragment Hap-B_2, can rescue amber mutations within gene VIII[8]. These observations suggest that the promoter which is responsible for the expression of gene VIII on fragment Hap-B_2 is located on this '300-fragment'. This, together with the observation that the 8S mRNA is only 360 nucleotides long, suggests that the central termination site for transcription is located somewhere within the '500-fragment' (fragment Hap-B_2·Hae-C).

In order to localize this termination site more precisely, we have studied in vitro RNA synthesis both under direction of the '300-fragment' (Hap-B_2·Hae-B) as well as under the direction of the '500-fragment' (Hap-B_2·Hae-C). As shown in Figure 3b, fragment Hap-B_2·Hae-B directs the synthesis of an RNA species which is approximately 210 nucleotides long. This size is in fact 150 nucleotides shorter than the size of the 8S RNA transcribed from fragment Hap-B_2 (Fig. 3a).Under the reaction conditions used, no RNA synthesis could be demonstrated when the '500-fragment' (Hap-B_2·Hae-C) was added to the in vitro RNA synthesizing system. Based on the assumption that termination of RNA synthesis on the '300-fragment' occurs immediately proximal to its 3'-terminal end (cf. Heyden et al.[20]), we conclude that the biologically most important signals which determine the size

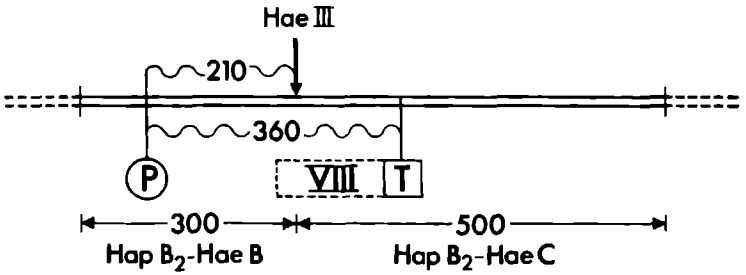


Fig. 5 : Schematic diagram of a portion of the M13 genome. P and T refer to the respective positions of the promoter and termination site on fragment Hap-B_2. The Hae III enzyme cleavage site on this fragment is indicated. The wavy lines refer to the DNA products transcribed either from fragment Hap-B_2 or from the '300-fragment'. The sizes of the restriction fragments and of the RNA transcripts are indicated by arabic numerals. The roman numeral refers to the location of gene VIII on fragment Hap-B_2.

of the 8S mRNA (promoter and termination site) are located, respectively, at a distance of approximately 90 and 450 nucleotides from the left-hand side of fragment Hap-B_2 (Fig. 5).

The latter conclusion could firmly be supported by *in vitro* transcription studies on restriction fragment Alu-D (Fig. 2), a fragment which encompasses 380 base pairs of the left-hand side of fragment Hap-B_2[9]. In the latter case only the synthesis of an RNA species with a size of 300 nucleotides could be demonstrated, *i.e.* 60 nucleotides shorter than that of the 8S RNA (Fig. 3c).

It should be noted that neither the '300-fragment' nor the fragment Alu-D are able to direct the *in vitro* synthesis of the protein encoded by gene VIII (Konings *et al.*, unpublished results). These observations suggest that only part of gene VIII is located on these restriction fragments and, therefore, the nucleotide sequence which codes for gene VIII protein is located proximal to the 3'- rather than the 5'-end of the 8S mRNA (Fig. 5).

A particular observation has to be emphasized. The size of the 8S mRNA (360 nucleotides) is about twice as large as the size of the mRNA (about 180 nucleotides) required to code for gene VIII protein (mol wt 5,800)[14]. Two hypotheses could account for this observation :

1) The 8S mRNA contains a stretch of RNA at its 5'-terminal end which does not code for protein. In this connection it is interesting to note that Sugimoto *et al.*[15] have found a long stretch of RNA which does not code for protein at the 5'-end of the RNA initiated at the phage fd promoter which is located on restriction fragment Hap-C (map position 0.94; Fig. 2).
2) The 8S mRNA contains the coding information for both the proteins encoded by gene VIII as well as by gene VII, a gene which is located on the genetic map immediately proximal to the '5'-end' of gene VIII[8] (Fig. 2).

At the present time there is no experimental data which would either refute or validate the first hypothesis. We favour, however, the latter hypothesis since genetic studies have indicated that all of the amber mutations within gene VII, tested so far, are rescued by fragment Hap-B_2[8] (Van den Hondel *et al.*, unpublished results). Furthermore, indirect observations have indicated that the size of gene VII must be small i.e. 350 nucleotides or less[8,18]. Although one should expect, in this case, that the 8S RNA should also direct the *in vitro* synthesis of gene VII protein, the latter is absent or present at an undetectable level (Fig. 4c). As a matter of fact, and for reasons still unknown, the synthesis of gene VII protein has not yet been detected *in vivo*[16,17] as well as *in vitro*[13,14,18].

Nucleotide sequence analysis of the 8S RNA transcript might soon afford additional information which may test the validity of either alternative. In addition, the exact location of the terminator, as reported in this study, together with the availability of restriction fragments containing this site, will generate further studies in order to elucidate the nucleotide sequences involved in factor-independent termination.

REFERENCES

1. Okamoto, T., Sugiura, M. and Takanami, M. (1969), *J. Mol. Biol.*, 45, 101-111.
2. Edens, L., Wezenbeek van, P., Konings, R.N.H. and Schoenmakers, J.G.G., *in preparation.*
3. Sugiura, M., Okamoto, T. and Takanami, M. (1969), *J. Mol. Biol.*, 43, 299-315.
4. Takanami, M., Okamoto, T. and Sugiura, M. (1970), *Cold Spring Harbor Symp. Quant. Biol.*, 35, 179-187.
5. Takanami, M. and Okamoto, T. (1974) in: *Control of Transcription*, chapter 12, pp. 145-156, Plenum Press, New York.
6. Pratt, D., Tzagoloff, H. and Beaudoin, J. (1969), *Virology*, 39, 42-53.
7. Hondel van den, C.A. and Schoenmakers, J.G.G. (1975), *Eur. J. Biochem.*, 53, 547-558.
8. Hondel van den, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G., *Eur. J. Biochem.*, (1975), 53, 559-567.
9. Hondel van den, C.A., Pennings, L. and Schoenmakers, J.G.G., *in preparation.*
10. Seeburg, P.H. and Schaller, H. (1975), *J. Mol. Biol.*, 92, 261-277.
11. Konings, R.N.H., Ward, R., Francke, B. and Hofschneider, P.H., (1970), *Nature*, 226, 604-607.
12. Staynov, D.Z., Pinder, J.C. and Gratzer, W.B., (1972), *Nature New Biology*, 108-110.
13. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975), *J. Virol.*, 15, 570-584.
14. Hondel van den, C.A., Konings, R.N.H. and Schoenmakers, J.G., (1975), *Virology*, in press.
15. Sugimoto, K., Okamoto, T., Sugisaki, H. and Takanami, M. (1975), *Nature*, 253, 410-414,
16. Henry, T.J. and Pratt, D. (1969), *Proc. Natl. Acad. Sci. U.S.A.*, 62, 800-807.
17. Lin, N.S.-C. and Pratt, D. (1974), *Virology*, 61, 334-342.
18. Model, P. and Zinder, N.D. (1974), *J. Mol. Biol.*, 83, 231-251.
19. Konings, R.N.H., Janssen, J., Cuypers, T. and Schoenmakers, J.G.G. (1973), *J. Virol.*, 12, 1466-1472.

20. Heyden, B., Nüsslein, C. and Schaller, H. (1975), *Eur. J. Biochem.*, 55, 147-155.
21. Konings, R.N.H., Stevens, M. and Schoenmakers, J.G.G., *in preparation.*
22. Franklin, R.M. (1966), *Proc. Natl. Acad. Sci. U.S.A.*, 55, 1504-1511.

CHAPTER IV

MAPPING OF PROMOTER SITES ON THE GENOME OF BACTERIOPHAGE M13

MAPPING OF PROMOTER SITES ON THE GENOME OF BACTERIOPHAGE M13

ABSTRACT

With the aid of transcription studies on restriction fragments of bacteriophage M13 DNA it has been demonstrated that at least eight promoter sites are located on the M13 genome. Five of these promoters initiate the synthesis of RNA chains which contain at their 5'-terminal end pppG (G-promoters), while the other three promoters initiate RNA chains which start exclusively with pppA (A-promoters). The positions of these promoter sites on the physical map are: 0.82 ($G_{0.82}$), 0.88 ($G_{0.88}$), 0.94 ($G_{0.94}$), 0.01 ($G_{0.01}$), 0.08 ($G_{0.08}$), 0.36 ($A_{0.36}$), 0.51 ($A_{0.51}$) and 0.56 ($A_{0.56}$). The G-promoters were found to be clustered within a distance of one-third of the genome length from the central termination site for transcription (map position 0.77). The A-promoters, however, were found at greater distances from this termination signal. Based upon the incorporation of (γ-^{32}P) ATP or (γ-^{32}P)GTP, the capacity of these promoters to initiate the synthesis of RNA chains varies. The strongest G-promoters are $G_{0.82}$, $G_{0.94}$ and $G_{0.08}$ and the strongest A-promoter is $A_{0.36}$.

As judged from their position on the genetic map, it is postulated that two promoters, namely $G_{0.94}$ and $G_{0.01}$, are located within gene II. The other promoters are most probably located immediately in front of the gene VIII/VII boundary ($G_{0.82}$), and immediately in front of gene V ($G_{0.88}$), gene II ($G_{0.08}$), gene IV ($A_{0.36}$), gene I ($A_{0.51}$) and gene VI ($A_{0.56}$). No evidence has been obtained so far for the existance of a promoter immediately in front of gene III.

INTRODUCTION

The genome of the small filamentous bacteriophage M13 consists of a circular, single-stranded DNA ($M_r \simeq 2 \times 10^6$) which, upon infection, is converted into a double-stranded, circular replicative form (RF) molecule (for a recent review, see(1)). During M13 growth only the nonviral strand of the RF molecule is transcribed by its host RNA polymerase (2-4). By convention transcription proceeds counterclockwise on the genetic map and Roman numerals are used to designate the eight M13 genes (5-7) (Fig. 1).

In a previous communication (8) we have presented evidence that transcription *in vitro* of M13 replicative form I results in the formation of at least seven discrete RNA species ranging in size from about 8S ($M_r \simeq 1.1 \times 10^5$) up to 26S ($M_r \simeq 17 \times 10^5$). The two largest RNA species (23S and 26S) are initiated with pppA (A-start RNA) while the other five (8S, 11S, 14S, 17S and 19S) are

initiated with pppG (G-start RNA) (8). Translation of these individual RNA species in a protein synthesizing system revealed that they all share the coding information for the major capsid protein encoded by gene VIII (our unpublished results). From these results it was inferred that the synthesis of all RNA species is terminated at a unique termination signal which is located immediately distal to the 3'-end of gene VIII. The latter conclusion has recently been substantiated by transcription studies *in vitro* on a DNA restriction fragment (fragment HapII-B_2; Fig. 1) which previously has been shown to direct the synthesis of gene VIII protein in a DNA-dependent protein synthesizing system (8,10).

The synthesis *in vitro* of a discrete number of RNA species implies that an equivalent number of RNA initiation sites (promoter sites) are located on the double-stranded replicative form molecule. The approximate positions of three of these promoters have already been mapped with the aid of coupled transcription and translation studies *in vitro* (7,10). However, in order to map these and the other promoters more precisely we have studied RNA synthesis under the direction of various restriction fragments of M13 RF molecules. From the results of such experiments it became evident that only those fragments gave rise to the synthesis of a discrete RNA product which were presumed to contain a promoter site on the basis of their location on the physical map. The length of the RNA chain formed then was determining for the position of the promoter on this fragment.

It could be demonstrated that the number of promoter sites detected on restriction fragments and the ribonucleoside triphosphates used for initiation are consistent with the number and starting nucleotide sequences of the RNA chains synthesized on the intact M13 genome.

MATERIALS AND METHODS

Enzymes

Escherichia coli RNA polymerase holoenzyme was a generous gift from Dr. R. Schilperoort (University of Leiden). The restriction endonucleases endoR · HindII, endoR · HapII and endoR· HaeIII have been isolated as described previously (14-16).

Replicative form I DNA and restriction fragments

The method for the preparative isolation and purification of M13 replicative form I DNA has been described (14,16). Digestion of it with the restriction endonucleases was carried out under the conditions previously described (5). After digestion, the fragments were separated by electrophoresis on a 3% discontinuous polyacrylamide slab gel (14) and further purified as described by van den Hondel *et al.* (5).

Full-length linear RF III molecules, produced by cleavage of M13 RF I with endoR' *HindII*, was isolated by sucrose density gradient centrifugation (11).

RNA synthesis *in vitro*

RNA synthesis *in vitro* was performed in a standard reaction mixture (0.1 ml) which contained: 4 µmol Tris-HCl (pH 7.9), 15 µmol KCl, 0.8 µmol $MgCl_2$, 0.1 µmol dithiothreitol, 0.01 µmol EDTA, 0.1% Tween-80, 25 µg bovine serum albumin, 0.2 pmol M13 DNA restriction fragment and 8.0 pmol *E.coli* RNA polymerase holoenzyme. After a preincubation period of 5 min at 37°C, ribonucleoside triphosphates were added to a final concentration of 80 µM, except for the (γ-^{32}P)-labeled triphosphate whose concentration was 6 µM. Initiation was allowed to proceed for 5 min at 37°C after which time the ATP or GTP concentration was adjusted to 80 µM and rifampicin was added (final concentration 25 µg/ml). RNA synthesis was then continued for 5 min at 37°C.

M13 RF I or full-length RF III molecules were transcribed under similar reaction conditions with the exception that the molar ratio of RNA polymerase to DNA template was 10.

After RNA synthesis, 10 µg of carrier tRNA was added and the reaction mixtures were extracted with phenol. The aqueous layer was passed through a Sephadex G-50 column (10x1cm) and the RNA was precipitated twice with 2.5 vol. of ethanol at -20°C. The RNA was collected by high-speed centrifugation (15 min at 150.000xg) and dissolved in 15 µl of deionized formamide.

Polyacrylamide gel electrophoresis

The RNA species synthesized *in vitro* were analysed on vertical slab gels (16x14x0.2cm), containing 3.8% polyacrylamide in 98% formamide (13). After electrophoresis for 4 h at 30 mA, the wet gel slab was exposed to X-ray film (Kodak RP/R54). The approximate lengths of the RNA species synthesized *in vitro* were estimated from their relative electrophoretic mobilities. As electrophoretic markers *E.coli* ribosomal RNAs and denatured restriction fragments of M13 RF were used. The chain lengths for the various RNAs was assumed to be 3000 nucleotides for 23S, 1500 nucleotides for 16S and 120 nucleotides for 5S ribosomal RNA (17). For the lengths of the restriction fragments the data given by van den Hondel and Schoenmakers (14) were used. On the gel system used no detectable differences were observed between the relative electrophoretic mobilities of the ribosomal RNAs and the restriction fragments of identical length.

RNA-DNA hybridization

The method for RNA-DNA hybridization was similar to the method of Gillespie and Spiegelman (18) as modified by Petterson *et al.* (19).

Nomenclature of promoter sites

Promoters which initiate the synthesis of RNA chains which start with pppG are denoted by G, while those which initiate the synthesis of RNA which starts with pppA are denoted by A. The position of each promoter is given in map units and is indicated by a suffix corresponding to the position of the promoter on the physical map (i.e. $G_{0.82}$ means a promoter which is located at 0.82 map units on the physical map and the RNA of which is started with pppG).

ABBREVIATIONS used: RF, double-stranded replicative form DNA; RF I, RF in which both strands are covalently closed; RF III, linear RF of full genome-length; HapII, HaeIII and HindII refer to the DNA fragments produced by the restriction endonucleases from H.aphirophilus (endoR·HapII), H.aegyptius (endoR HaeIII) and H.influenzae serotype d(endoR·HindII).

Note. The direction of transcription proceeds counterclockwise on the genetic map. This is also the 5' to 3' polarity of the restriction fragments.

RESULTS

We reported previously (8) that transcription of circular M13 replicative form I by purified E.coli RNA polymerase holoenzyme results in the formation of at least two large RNA species (23S and 26S) which are initiated with pppA and five species (8S, 11S, 14S, 17S and 19S) which are initiated with pppG. From their relative electrophoretic mobilities on the polyacrylamide gel one can calculate that these RNA species range in size from approximately 360 nucleotides for the smallest (8S) up to about 5000 nucleotides for the largest (26S) RNA product. Also the frequency of synthesis varies among the individual RNA species (Table 1). If it is taken into account that all these RNA species are terminated at the same site and that this central termination site is located immediately distal to gene VIII (8), then the approximate positions of initiation regions on the genetic and physical map can be determined. The positions estimated from transcription data on intact M13 RF molecules are presented in Table 1. Due to the inaccuracy of estimating sizes of RNA from their relative electrophoretic mobilities, especially in the high-molecular weight region, the positions given are only fairly accurate for sites which initiate the relatively small RNA chains. In order to verify the existence of such promoter regions and to map these promoter sites more accurately we studied RNA synthesis under the direction of restriction fragments of M13 RF. The fragments used were the endoR· HaeIII, the endoR· HapII and the endoR·HindII fragments, whose physical orders are illustrated in Fig. 1.

As shown in Table 2, RNA synthesized on restriction fragments with RNA polymerase holoenzyme did not hybridize to single-

RNA species	Promoter	Position of promoter as deduced from the length of RNA transcribed from			Estimated length of RNA chain (nucleotides)	Relative strength of promoter
		RF I	RF III	restriction fragments		
8S	$G_{0.82}$	0.80	0.80	0.82	360	0.50-0.80
11S	$G_{0.88}$	0.90	0.90	0.88	700	0.05-0.15
14S	$G_{0.94}$	0.95	0.95	0.94	1200	1.00
17S	$G_{0.01}$	0.00	0.00	0.01	1600	0.10-0.15
19S	$G_{0.08}$	0.10	0.10	0.08	2000	0.60-0.90
23S	$A_{0.36}$	0.30	0.35	0.36	3900	1.00
26S	$A_{0.51}$	n.d.	0.50	0.51	4800	0.15-0.30
26S	$A_{0.56}$	n.d.	0.50	0.56	5100	

Table 1: Relative strengths and positions of the promoters on the M13 genome. The RNA species synthesized in vitro under the direction of M13 RF have been designated 8S, 11S, 14S, etc.(8). The sizes of these RNA chains were approximated from their relative electrophoretic mobilities on formamide slab gels with E.coli ribosomal RNA as markers. The lengths of the RNA species synthesized on restriction fragments of M13 RF were estimated similarly using E.coli ribosomal RNA and denatured restriction fragments as markers. The position of a promoter on a restriction fragment was calculated from the length of the RNA chain synthesized on this fragment, on the assumption that synthesis has ceased at its 3'end. The length of the RNA chains synthesized on M13 RF finally has been calculated from the distance between the central termination site (map position 0.77 (8)) and the promoter site as determined by transcription of the various restriction fragments. The strength of the various promoters was estimated from the ^{32}P content of the RNA chains initiated at the respective promoter sites, using either (γ-^{32}P)ATP or (γ-^{32}P)GTP as the sole radioactive precursors during transcription. The relative strength of the G promoters and A promoters were evaluated and expressed as a fraction of the strength of promoter $G_{0.94}$ and $A_{0.36}$, respectively. n.d.= not determined.

stranded phage DNA but did find complements with denatured M13 RF. This indicates that under the transcription conditions used the RNA is predominantly transcribed from the minus strand of the restriction fragments, i.e. the strand which is also used as a template during the transcription *in vivo* and *in vitro* of the intact M13 genome (2,3). Only a limited number of restriction fragments of M13 RF were able to direct the synthesis of a unique RNA product, as will be shown in the following paragraphs: namely those fragments which were presumed to contain a promoter on the basis of their location on the physical map (Table 1). These combined results indicate that proper strand selection and the specificity of RNA initiations is still retained after cleavage of M13 RF molecules into restriction fragments. The efficiency of initiation, however, appeared to be influenced by restriction endonuclease digestion.

Replicative form used for transcription	Amounts of (^{32}P) RNA hybridized to		
	M13 RF	M13 viral DNA	T7 DNA
	%		
Untreated M13 RF	65	0.7	0.4
EndoR· *Hap*II-treated RF	51	2.2	0.6
EndoR· *Hae*III-treated RF	53	1.3	0.4

Table 2: Hybridisation of bacteriophage M13 RNA synthesized *in vitro*. ^{32}P-labeled RNA was synthesized *in vitro* either on M13 RF I or on RF which has been digested with the appropriate restriction endonucleases. About 0.04 µg RNA (5.5×10^4 counts · min^{-1}) was incubated for 20 h at 37°C with either heat-denatured M13 RF (2 µg), M13 ssDNA (1 µg), or denatured T7 DNA (2 µg) which were immobilized on nitrocellulose filters. The hybridisation buffer consisted of 50% formamide, 0.6M NACl, 0.05 M Tris-HCl, pH 7.4, 10 mM EDTA and 0.1% sodium dodecylsulfate. Non-hybridized RNA was removed from the filters by digestion with pancreatic ribonuclease and ribonuclease T1. Subsequently, the filters were washed several times with 0.03 M NaCl, 0.03 M sodium citrate buffer, dried, and counted in PPO/POPOP/toluene scintillation fluid.

Location of A-promoters

From their length in nucleotides it is to be expected that the synthesis of the ATP-initiated RNA chains starts at the map positions 0.3 (23S) and 0.5 (26S), respectively (Table 1 and Fig. 1). Fragments which encompass these sites and, hence, are presumed to contain these A-promoters regions are HapII-A, HapII-B_1, HaeIII-A and RF III. The transcriptional results obtained with these fragments are described in the following sections.

Fragment HapII-A

Genetic analysis and coupled transcription and translation studies *in vitro* have demonstrated that the complete gene IV is located on fragment HapII-A (≈1530 base pairs long (5,7)). Since the size of gene IV is almost equivalent to the size of fragment HapII-A it was concluded that a promoter is located proximal to the 5'end of this restriction fragment.

As shown in Fig. 2h, transcription of fragment HapII-A in the presence of (γ-^{32}P)-labeled ATP and subsequent analysis of the products on formamide polyacrylamide slab gels revealed that this fragment directs the synthesis of a single RNA species which is approximately 1400 nucleotides long. Similar to fragment Hap II-A, this RNA species directs in a protein synthesizing system the synthesis of the polypeptide encoded by gene IV (7) (and unpublished results). These observations suggest, therefore, that this RNA species is initiated at the same promoter on fragment HapII-A as is the RNA transcribed from this fragment in the DNA-dependent protein synthesizing system (7,10). This A-promoter, designated $A_{0.36}$, therefore must be located at or within a distance of approximately 150 nucleotides from the 5'end of this restriction fragment (Fig. 1).

Fragment HapII-B_1

With the aid of marker rescue experiments we have demonstrated that restriction fragment HapII-B_1 (≈800 base pairs long) contains genetic markers for both gene VI and I and very likely

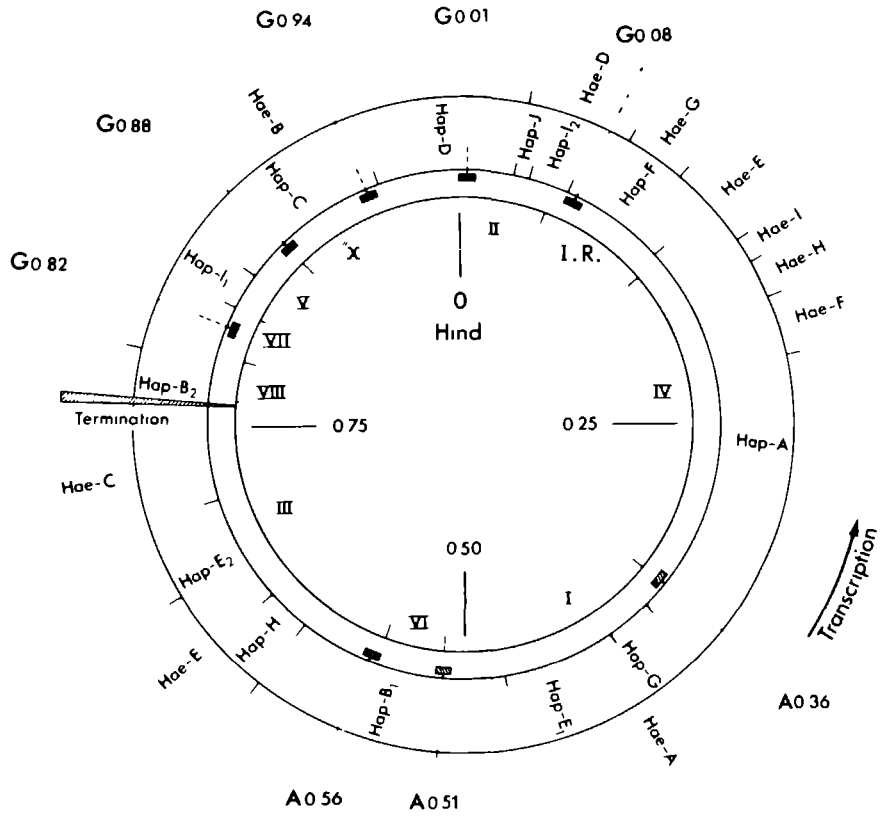


Fig. 1: Localization of promoter sites on the genetic map and the physical maps of bacteriophage M13 DNA. The inner circle represents the genetic map (5). The middle and outer circles show the locations of the endoR· HapII and the endoR· HaeIII fragments (14). The arrow indicates the cleavage site of M13 RF by endoR· HindII (14). The direction of transcription is counterclockwise around the genetic map (7). The positions of the promoters discussed in this paper are indicated: G promoters with black bars, and A promoters with hatched bars. The position of the central termination site for transcription (8) is indicated: 'X' refers to the C-terminal part of gene II in which an internal start of RNA synthesis is located. This RNA directs the synthesis of protein X (10,15) (and Edens,L., Konings, R.N.H. and Schoenmakers, J.G.G., unpublished results). I.R. refers to the intergenic region between gene II and gene IV (15,32), which contains the replication origin for parental M13 RF (23).

also for gene III (5, 15). The order of these genes on the genetic map is 5'-III-VI-I-3' (Fig. 1).

Contrary to expections, transcription of fragment HapII-B_1 with RNA polymerase holoenzyme resulted in the formation of two, instead of one, RNA species, the sizes of which are approximately 500 and 200 nucleotides long (Fig. 2j). Comparison of the bands on the autoradiograph indicated that they differ in intensity. This suggests that these RNA species are transcribed from promoters which differ in their affinity for RNA polymerase holoenzyme.

Genetic analysis as well as protein synthesis studies in vitro (5,7,15) have indicated that the size of gene VI must be small and, furthermore, that the N-terminal ends of both genes VI and I are located proximal rather than distal to the 3'end of fragment HapII-B. Since gene VI and gene I are adjacent to each other on the genetic map it is attractive to postulate that the promoters of the RNA species transcribed from fragment HapII-B_1 are located immediately in front of the N-terminal ends of genes VI ($A_{0.56}$) and I ($A_{0.51}$), i.e. at 500 and 200 nucleotides from the 3'end of fragment HapII-B_1 (Fig. 1).

Fragment HaeIII-A

In earlier reports (5,7,10) evidence was given that fragment HaeIII-A encompasses the C-terminal end of gene III, the complete genes VI and I and about 1000 nucleotides of the N-terminal end of gene IV (Fig. 1). As shown on the cleavage maps of M13 RF the complete nucleotide sequences of fragments HapII-B_1, HapII-E_1 and HapII-G, and about 1000 nucleotides of the 5' end of fragment HapII-A are contained on this large restriction fragment (Fig. 1). Transcription of fragment HaeIII-A in the presence of (γ-^{32}P)ATP results in the formation of two major RNA species which are about 2000 and 850 nucleotides long (Fig. 2i), and a minor one, the size of which is approximately 1750 nucleotides and which is recognized as a faint band just beneath the largest RNA product on the formamide gel (Fig. 2i).

In the previous sections we have demonstrated that the RNA species transcribed from fragment HapII-B_1 are approximately

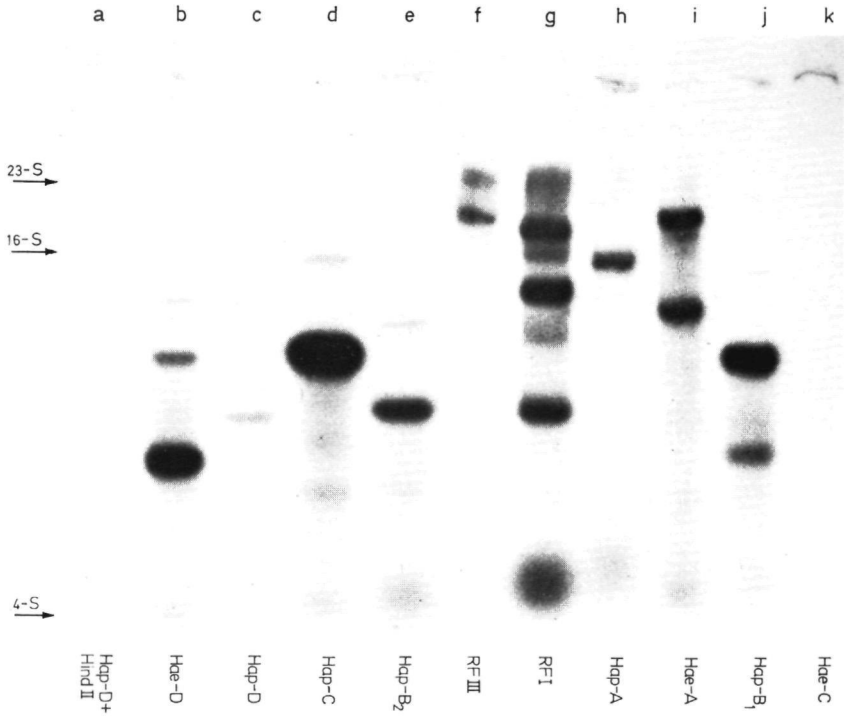


Fig. 2: Autoradiograph of the RNA species synthesized under the direction of M13 RF I or under the direction of various fragments of M13 RF. RNA synthesis was carried out in the presence of (γ-^{32}P)ATP or (γ-^{32}P)GTP and was performed as described under Methods. The RNA products were analysed on vertical slab gels (16x14x0.2cm) containing 3.8% polyacrylamide in 98% formamide (13). The arrows indicate the positions of migration of the ribosomal RNA species isolated from E.coli cells. The RNA products were synthesized under the direction of: (a) HapII-D, which was split with endoR·HindII, (b) HaeIII-D, (c) HapII-D, (d) HapII-C, (e) HapII-B_2, (f) RF III, i.e. RF which was split with endoR·HindII, (g) RFI, (h) HapII-A, (i) HaeIII-A, (j) HapII-B_1, (k) HaeIII-C. The synthesis of the RNA products (a-e,g) were carried out in the presence of (γ-^{32}P)GTP, whereas the others (h-k,f) were carried out in the presence of (γ-^{32}P)ATP.

500 and 200 nucleotides long. Fragment HapII-B_1, however, is about 1650 nucleotides shorter at its 3'end than fragment HaeIII-A (5). This evidence, therefore, strongly suggests that the two largest RNA species transcribed from fragment HaeIII-A are initiated at the same promoters ($A_{0.56}$ and $A_{0.51}$, Fig. 1) as are the RNA species transcribed from fragment HapII-B_1. Moreover, our finding that the synthesis of the 1750 nucleotides-long RNA products is much lower than the 2000 nucleotides-long transcript is consistent with the observed character of these promoters of which the $A_{0.51}$ promoter has the weakest affinity for RNA polymerase.

In the previous section we also demonstrated that the $A_{0.36}$ promoter is located at or within the first 150 nucleotides of fragment HapII-A. Since fragment HaeIII-A encompasses the first 1000 nucleotides of fragment HapII-A, it is concluded that approximately 850 nucleotides of the N-terminal end of gene IV are located on this HaeIII-A fragment. The observation that the smallest RNA species transcribed from fragment HaeIII-A is almost as long as this N-terminal end, i.e. about 850 nucleotides, strongly suggests that this RNA species is initiated at the promoter ($A_{0.36}$) which is located in front of gene IV. The latter conclusion is completely consistent with our previous results of coupled transcription and translation studies which showed that fragment HaeIII-A is only capable of synthesizing the N-terminal part of gene IV protein (7,10).

RF III

Cleavage of circular M13 RF I with the restriction endonuclease R'HindII results in the formation of a linear double-stranded molecule (RF III) of genome length (11). The single cleavage site for this restriction enzyme has been defined on the physical map at zero point position (Fig. 1) (14,16).

As is shown in Fig. 2f and 3b, transcription of RF III in the presence of (γ-^{32}P)ATP, followed by electrophoretic analysis on formamide slab gels, gives rise to the formation of two ATP-initiated RNA chains, the sizes of which are approximately 2100 and 3200 nucleotides. If it is assumed that the synthesis

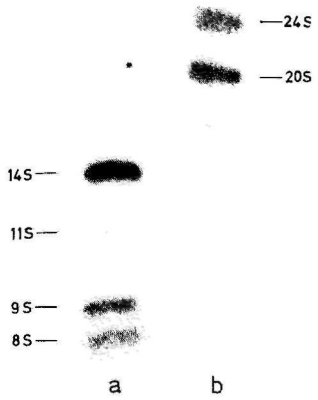


Fig. 3: Autoradiograph of the RNA species synthesized under the direction of M13 RF III. M13 RF III was obtained by cleavage of RF I with endoR· HindII (10). The conditions for RNA synthesis in vitro and the analysis of the RNA products on 3.8% polyacrylamide slab gels in 98% formamide have been described under Methods. (a) RNA products synthesized on M13 RF III in the presence of (γ-^{32}P)GTP. (b) RNA products synthesized on M13 RF III in the presence of (γ-^{32}P)ATP.

of these RNA chains has ceased at the 3'-terminal end of the linear RF molecule (cf. (20)), their starting positions should be located at about 0.3 and 0.5 map units (Table 1 and Fig. 1). These positions are almost identical to the positions of the $A_{0.36}$ and $A_{0.56}$ (or $A_{0.51}$) promoters as deduced from the transcription data on the fragments HaeIII-A, HapII-A and HapII-B_1 (see above).

It cannot be ascertained whether the largest RNA species (3200 nucleotides long) is a homogeneous transcriptional product of the strong promoter $A_{0.56}$ or consists of a mixture of chains, which deviate only slightly in lenght and which originate from both the $A_{0.56}$ promoter and the (weak) $A_{0.51}$ promoter.

Location of G-promoters

Previously we reported that five promoters are present on the M13 genome which give rise to the synthesis of pppG-initiated RNA chains (8). The relative strenghts of these G-promoters, as deduced from the incorporation values (γ-^{32}P)GTP into RNA chains, and their approximate map positions, as deduced from the length of RNA chains synthesized on intact RF I, are given in Table 1.

In order to map these G-promoters more accurately, the same strategy as used above for the mapping of the A-promoters was followed.

Fragment HaeIII-D

Restriction fragment HaeIII-D, which is approximately 310 base pairs long (14), was recently shown to encompass a part of the intergenic space between gene II and gene IV and a part of the N-terminal end of gene II (15).

Transcription of this fragment results in the formation of one major RNA product of about 250 nucleotides long and several minor RNA species (Fig. 2b). Interestingly, the minor species, the sizes of which are about 600 and 900 nucleotides, are one and more fragment-length longer than the major RNA transcript. These observations strongly suggest that termination of transcription does not always occur at the 3'-terminal end of this particular fragment (cf. (20-22)) but that the RNA polymerase infrequently is able to switch transcription at this end from one strand to the other. The observation that the decrease in intensity of the RNA bands parallels their increase in size is consistent with the latter hypothesis. In this respect it has to be emphasized that fragment HaeIII-D is the sole fragment we have isolated so far which demonstrates this template switching so clearly. Since the replication origin for complementary strand synthesis of progeny M13 RF is located near this fragment (23), it is possible that this fragment has a particular secondary and tertiary structure. This in turn may explain why transcription is not always terminated at the 3'end of this restriction fragment.

Our observation that the N-terminal end of gene II is located on fragment HaeIII-D (15) together with the observation that the RNA initiation site is located at a distance of approximately 250 nucleotides from its 3'end, strongly suggests that this $G_{0.08}$ promoter is located in the intergenic region between gene II and gene IV at a position which is immediately proximal to the N-terminal end of gene II (Fig. 1).

Fragment HapII-D

As is shown in Fig. 1, restriction fragment HapII-D (≈560 base pairs long) encompasses almost completely the first half of gene II (5,7,15). It does not contain, however, the N-terminal of this gene. Transcription of fragment HapII-D results in the weak synthesis of an RNA species which is about 350 nucleotides long (Fig. 2c). This observation suggests that a promoter, namely promoter $G_{0.01}$ (Fig. 1), is located on this fragment at approximately 200 base pairs distal to the 5'end.

Cleavage of restriction fragment HapII-D with the restriction endonuclease R· HindII results in the formation of two new DNA fragments which are 240 and 300 base pairs long (14). The 5'end of fragment HapII-D is located on the smallest of these two fragments (Fig. 1). Transcription of either fragment did not result in the formation of detectable amounts of RNA (Fig. 2a). These observations may be explained by the following alternatives: either (a) the biological activity of the promoter site on fragment HapII-D is completely destroyed due to cleavage of this fragment by the restriction of endonuclease R· HindII or (b) the promoter-containing DNA fragment which remains after cleavage of HapII-D with endoR· HindII gives rise to the synthesis of such a small RNA species that its presence on the polyacrylamide gel is difficult to detect. Independent of the correct explanation, the results are in any case in full agreement with the conclusion that the $G_{0.01}$ promoter on HapII-D is located at about 200 nucleotides from the 5'end of this fragment.

Fragment HapII-C

Genetic studies as well as protein synthesis studies *in vitro*, have recently indicated that this fragment, which is about 650 base pairs long, constitutes the C-terminal part of gene II and the N-terminal part of gene V (5,10,15). The order of these genes on the genetic map is 5'-II-V-3' (Fig. 1). Transcription of fragment HapII-C results in the formation of one major and one minor RNA product, the sizes of which are 570

and 180 nucleotides, respectively (Fig. 2d). The promoters which initiate these pppG-containing RNA chains, therefore, have to be positioned at 0.94 and 0.88 map units (Fig. 1). As judged from the intensity of the bands on the autoradiograph, it is evident that the promoters $G_{0.94}$ and $G_{0.88}$ differ significantly in their affinity for RNA polymerase. This difference in not only noted with fragment HapII-C but is also apparent when intact RF I is used as a template for transcription (cf. Table 1).

Infrequently, the weak synthesis of an RNA species (≈1300 nucleotides), which was about twice as long as the fragment template, could be observed. It is probable that the latter product is initiated at the strong $G_{0.94}$ promoter, which is located proximal to the 5'end of fragment HapII-C. Elongation then proceeds beyond the 3'end of this fragment, most probably by switching transcription from one strand to the other.

DNA-dependent protein synthesis studies *in vitro* have demonstrated that fragment HapII-C directs the synthesis of protein X, a polypeptide which is about 130 amino acid residues long (5,7) and whose complete genetic information is located within the C-terminal end of gene II (15,33). From our present observation that the major RNA product of fragment HapII-C, which is initiated at the strong $G_{0.94}$ promoter, is large enough to code for this polypeptide, we infer that in front of this 'gene' the $G_{0.94}$ promoter is positioned.

Fragment HapII-B_2

Restriction fragment HapII-B_2 is approximately 800 base pairs long (Fig. 1). In a DNA-dependent cell-free system it directs the synthesis of the polypeptide encoded by gene VIII (10). Genetic analysis and transcription and translation studies *in vitro* have indicated that besides gene VIII also gene VII and the N-terminal part of gene III are located on this restriction fragment (5,7,8,10,15). The order of these genes on the genetic map is 5'-VII-VIII-III-3' (Fig. 1).

Previously, we have demonstrated that the promoter $G_{0.82}$ as well as the central termination site for transcription are

located on fragment HapII-B_2 (8). These regulatory sites, which mark the ends of 8S RNA (360 nucleotides), are located at 90 and 450 nucleotides, respectively, from the 5' end of this restriction fragment (Fig. 1).

As shown in Fig. 2e, fragment HapII-B_2 directs, besides the synthesis of the 8S RNA, also the weak synthesis of an RNA species which is about 720 nucleotides long. The 8S RNA product has previously been shown to code for the synthesis of gene VIII protein only (8,33). The length of the larger RNA product corresponds to the distance between the $G_{0.82}$ promoter and the 3' end of fragment HapII-B_2. This observation, together with the result that this large RNA is also able to direct the synthesis of gene VIII protein (unpublished results) suggests that this 720 nucleotides-long RNA is the product of a transcription process which is initiated at promoter $G_{0.82}$ but which is not terminated until the RNA polymerase has reached the 3'end of this restriction fragment. From these results and our unpublished data we infer that the RNA polymerase holenzyme leaks infrequently, for reasons still unknown, through this rho-independent termination signal which is located close to the center of fragment HapII-B_2.

RF III

Previously we demonstrated that transcription of RF I molecules in the presence of (γ-^{32}P)GTP results in the formation of five GTP-initiated RNA chains, the sizes of which correspond to 8S, 11S, 14S, 17S and 19S (8). This is demonstrated again in Fig. 2g. In contrast, transcription of RF III gives rise to only four GTP-initiated RNA chains of 360, 450, 700 and 1200 nucleotides (Fig. 3a). The former is identical to 8S RNA, the largest products correspond to 11S and 14S RNA, respectively. As compared to RF I, two RNA products, i.e. 17S and 19S RNA are missing but a shorter one, namely 9S RNA is present (Fig. 3a). These observations are consistent with our findings (above) that the endoR· HindII cleavage site on the M13 genome is located in or at least in the vicinity of the $G_{0.01}$ promoter. Transcription from the $G_{0.08}$ promoter then results in the formation of a

prematurely terminated RNA chain of about 450 nucleotides (9S RNA).

DISCUSSION

After infection of an Escherichia coli cell with bacteriophage M13 not all phage-specific proteins are synthesized in equimolar amounts (24). Some proteins, such as the 'DNA-unwinding protein' encoded by gene V and the major capsid protein encoded by gene VIII, are synthesized in much larger quantities than the other phage-specific proteins. The latter proteins are also the major products synthesized in M13 DNA-directed reactions in vitro (6,7,25). To explain these noticeable differences of gene expression several regulatory mechanisms have been proposed (25). The available experimental data (3,8,10,26,27,30) favour the model according to which the replicative form molecule is transcribed into several distinct polycistronic mRNAs which are initiated at different promoter sites (RNA initiation sites) but all are terminated at a unique termination signal which is located immediately after gene VIII (8,27).

There are several indications that such a regulatory mechanism of gene expression is not solely restricted to the filamentous coliphages but that it is also operative during the expression of the late genes of the phages T7 (28), T4 and λ (E. Young, B. Studier and M. Pearson, personal communications). Whether this mechanism of gene expression possesses a more universal character however, awaits further investigation.

The results presented in this paper provide strong evidence that at least eight promoter sites are located on the M13 genome. The map positions of these promoter sites are shown in Fig. 1. The RNA transcribed from three of these promoters ($A_{0.56}$, $A_{0.51}$ and $A_{0.36}$) are initiated with pppA (A-start RNA) while the RNAs transcribed from the other five promoters ($G_{0.08}$ through $G_{0.82}$) are initiated with pppG (G-start RNA). All of these RNA species share the coding information for the major capsid protein (our unpublished results) which support the observation that on the M13 genome a rho-independent termination of transcription occurs at a unique site which is located immediately

distal to gene VIII (8).

From the results of the experiments in which replicative form I molecules are transcribed in the presence of either (γ-^{32}P)ATP or (γ-^{32}P)GTP it can be concluded that the RNA chains are initiated with different frequencies (Table 1). Hence, both strong and weak promoters are located on the M13 genome. The strongest promoters are $G_{0.08}$, $G_{0.94}$, $G_{0.82}$ and $A_{0.36}$ (Fig. 1). Based upon the ^{32}P-radioactivity incorporated into each RNA chain, the remaining promoters ($G_{0.01}$, $G_{0.88}$, $A_{0.51}$ and $A_{0.56}$) revealed about 0.25-0.1-fold weaker affinity for RNA polymerase holoenzyme. Whether these differences in affinity reflect differences in nucleotide sequences between these promoters has yet to be established. From nucleotide sequence analysis of the strong promoters $G_{0.82}$ and $G_{0.94}$ of phage fd it is known that the longest sequence common to both promoters is only T-A-T-A-A-T (21, 22,29).

Recently, Seeburg and Schaller (30) have also studied promoter locations on the genome of M13 and its relatives fd and f1. Their method consisted of binding RNA polymerase to restriction fragments in the presence or absence of various triphosphate mixtures. The results obtained by this group cover to a large extent the results we have obtained since they also found promoter sites on the restriction fragments HapII-A, HapII-B_1, HapII-B_2, HapII-C, HapII-D and HapII-F. Unlike our observations they did not find two promoter sites on fragments HapII-B_1 and HapII-C. These discrepancies, however, are almost certainly due to the different techniques used for the location of promoter sites. Okamoto *et al.* (31) have studied the location of promoter sites on the fd genome with the aid of the filter-binding assay and, in addition, with the same technique used here. However, they detected only four promoter sites on the fd genome. Three of these promoters (G-promoters) were localized at the same map positions as we have found for the strong promoters $G_{0.82}$ (HapII-B_2), $G_{0.94}$ (HapII-C) and $G_{0.08}$ (HaeIII-D) on the M13 genome. The fourth promoter (A-promoter), however, was localized at map position 0.46 (Fig. 1). This position is completely different from the positions of the A-promoters found on the fd genome by Seeburg and Schaller (30) and on the M13 ge-

nome by our group. At present we have no reasonable explanation might be, though not very likely, that both groups are working with variants of the original fd strain.

Given the positions of the various M13 phage promoters on the physical map, it is very attractive to correlate their positions with the genetic map (Fig. 1). In previous communications (7,10) we have provided evidence for the existance of a promoter in front of gene IV, a promoter in front of gene VIII, and an RNA initiation site which was positioned somewhere in the middle of gene II. The promoters which accord with such positions are the $A_{0.36}$, the $G_{0.82}$ and the $G_{0.94}$ promoter, respectively. Several studies have demonstrated now that the deduced position of $G_{0.82}$ and $A_{0.36}$ in front of these genes is correct (8,32) (and our unpublished results). In the meantime, it has been shown that $G_{0.94}$ represents an internal start of RNA synthesis which is located within the C-terminal region of gene II and which leads to the synthesis of protein X (15). Moreover, transcription of several new restriction fragments of M13 RF and subsequent translation of the synthesized RNA products now have demonstrated unambiguously that promoter $G_{0.88}$ is positioned in front of gene V, that promoter $G_{0.08}$ is located in front of gene II and that $G_{0.01}$ represents a second but weak intragenic initiation site (our unpublished results).

The biological function of these intragenic initiation sites within gene II is not known. As far as protein X is concerned, we have no evidence at present whether the amino acid sequence of this protein is identical to the amino acid sequence of the C-terminal part of gene II protein.In other words, it is not known whether the mRNA coding for protein X is translated in the same reading frame as the mRNA coding for gene II protein. If not, one is left with the unique situation that one of the M13 genes, namely gene X, is enclosed within another gene, i.e. gene II.

Since we do not know the exact positions of the 5'ends of genes VI and I on the genetic map, it is impossible to define unambiguously which promoters are located in front of these genes. The experimental data presented in this study, however, favour the hypothesis that the promoters $A_{0.56}$ and $A_{0.51}$ are positioned

in front of these genes. Genetic studies have indicated that in vivo a polarity exists among genes III, VI and I in that the expression of genes VI and I is dependent on the expression of gene III. Our findings that promoters are located most probably in front of the N-terminal ends of genes I and VI suggests that both genes can be expressed independently of gene III. The biological implications of these discrepancies can not yet be understood.

Previously we have presented evidence that on the M13 genome termination of transcription occurs at a unique site which is located between the genes III and VIII(8). Furthermore, we have demonstrated (unpublished results) that the mixture of RNAs transcribed from M13 RF directs in a protein synthesizing system the synthesis of gene III protein. These results are in favour of a promoter in front of gene III. However, up to now we have failed to demonstrate the presence of such a promoter. Three hypothesis could account for these observations.

(a) A promoter is located in front of gene III but its capacity to bind RNA polymerase holoenzyme is destroyed by cleavage of the circular replicative form I molecules into linear DNA fragments.

(b) A promoter is located in front of gene III but its affinity for RNA polymerase is very weak resulting in the synthesis of very small amounts of high molecular weight RNA ($M_r \approx 2x\ 10^6$). Such small amounts of RNA might have escaped detection on the polyacrylamide gel system used.

(c) No promoter is located in front of gene III, but termination of transcription is not stringent. Hence, transcription of gene III is the result of a transcription process whereby RNA polymerase leaks infrequently through the central termination site located between the C-terminal end of gene VIII and the N-terminal end of gene III (Fig. 1). As a consequence, small amounts of high molecular weight transcripts of gene III are synthesized which are, by definition, heterogeneous in nature and therefore difficult to detect by electrophoretic analysis. Our observation that fragment HapII-B_2 (see Results) directs, besides 8S RNA, also the synthesis of a 720 nucleotides-long RNA which also carries the coding information of for gene VIII pro-

tein, can only be explained by such a 'read-through' transcription process and is, therefore, in agreement with the latter hypothesis.

The results obtained thus far are consistent with the multi-promoter single-terminator model suggested by Sugiura et al. (36). That the different size classes of phage M13 mRNA synthesized in vitro are indeed the result of such a "cascade" mechanism of transcription has recently been demonstrated unambiguously by transcription and subsequent translation studies of various M13 DNA restriction fragments (unpublished results).

The proposed cascade mechanism giving rise to RNA molecules with overlapping nucleotide sequences, may prove to be an important mechanism for regulating the amount of the various gene products of this phage. Particularly, the genes which are located proximal to the central termination site, i.e. gene VIII and gene V will be transcribed more frequently than the other genes resulting in the formation of larger amount of their encoded proteins. Our previous observations that the DNA-binding protein, encoded by gene V, and the major coat protein, encoded by gene VIII, are synthesized in vitro in much larger amounts than the other phage proteins (7,25) (cf. (6)) are in any case consistent with the proposed cascade mechanism.

Gene V protein and gene VIII protein are also the products which are most abundantly present in the infected cell (24). However, no direct evidence is available yet which demonstrates that the proposed model of transcription is also relevant for the control of M13 gene expression in vivo. Preliminary investigations indicate that in the infected cell several RNA species are present which range in size from 8S upto 30S and which share the coding information for the polypeptide encoded by gene VIII (unpublished results). Whether these RNA species are initiated at the same promoters and whether they are formed by cascade transcription of the M13 genome, is currently under investigation.

REFERENCES

1. Denhardt, D.T. (1975), *Crit. Rev. Microbiol.*, 4, 161-223.
2. Jacob, E. and Hofschneider, P.H. (1969), *J. Mol. Biol.*, 46, 359-363.
3. Okamoto, T., Sugiura, M. and Takanami, M. (1969), *J. Mol. Biol.*, 45, 101-111.
4. Jacob, E., Jaenisch, R. and Hofschneider, P.H. (1970), *Nature, (Lond.)*, 277, 59-60.
5. Van den Hondel, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975), *Eur. J. Biochem.*, 53, 559-567.
6. Model, P. and Zinder, N.D. (1974), *J. Mol. Biol.*, 83, 231-251.
7. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975, *J. Virol.*, 15, 570-584.
8. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975), *Nucleic Acid Res.*, 2, 1811-1820.
9. Reference deleted.
10. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975), *Virology*, 67, 487-497.
11. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1973), *Mol. Biol. Rep.*, 1, 41-45.
12. Konings, R.N.H., Jansen, J., Cuypers, T. and Schoenmakers, J.G.G. (1973), *J. Virol.*, 12, 1466-1472.
13. Staynov, D.Z., Pinder, J.C. and Gratzer, W.B. (1972), *Nat. New Biology*, 235, 108-110.
14. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1975), *Eur. J. Biochem.*, 53, 547-558.
15. Van den Hondel, C.A., Pennings, L. and Schoenmakers, J.G.G., (1976), *Eur. J. Biochem.*, 68, 55-70.
16. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1976), *J. Virol.*, 18, 1024-1039.
17. Fellner, P. (1974), in *Ribosomes* (M. Nomura, A. Tissières and P. Lengyl , eds.) pp. 169-191, Cold Spring Harbor Press, New York.
18. Gillespie, D. and Spiegelman, S. (1965), *J. Mol. Biol.*, 12, 829-842.

19. Petterson, U., Sambrook, J., Delius, H. and Tibbets, C. (1971), *Virology*, 59, 153-167.
20. Heyden, B., Nüsslein, C. and Schaller, H. (1975), *Eur. J. Biochem.*, 55, 147-155.
21. Schaller, H., Gray, C.P. and Herrmann, K. (1975), *Proc. Natl. Acad. Sci., U.S.A.*, 72, 737-741.
22. Sugimoto, K., Okamoto, T., Sugisaki, H. and Takanami, M., (1975), *Nature (Lond.)*, 253, 410-414.
23. Horiuchi, K. and Zinder, N.D. (1976), *Proc. Natl. Acad. Sci. U.S.A.*, 73, 2341-2345.
24. Henry, T.J. and Pratt, D. (1969), *Proc. Acad. Sci., U.S.A.*, 62, 800-807.
25. Konings, R.N.H. (1973), *FEBS Lett.*, 35, 155-160.
26. Takanami, M., Okamoto, T. and Sugiura, M. (1970), *Cold Spring Harbor Symp. Quant. Biol.*, 34, 179-187.
27. Chan, T., Model, P. and Zinder, N.D. (1975), *J. Mol. Biol.*, 99, 369-382.
28. Pachl, C.A. and Young, E.T. (1976), *Proc. Natl. Acad. Sci. U.S.A.*, 73, 312-316.
29. Takanami, M., Sugimoto, K., Sugisaki, H. and Okamoto, T. (1976), *Nature (Lond.)*, 260, 297-302.
30. Seeburg, P.H. and Schaller, H. (1975), *J. Mol. Biol.*, 92, 261-277.
31. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, M. (1975), *J. Mol. Biol.*, 95, 33-44.
32. Vovis, G.F., Horiuchi, K. and Zinder, N.D. (1975), *J. Virol.* 16, 674-684.
33. Pratt, D., Tzagoloff, H., Erdahl, W.S. and Henry, T.J. (1967), in *Molecular Biology of Viruses* (J.S. Colter and W. Parachych, eds.)pp. 219-238, Academic Press, New York.
34. Lyons, L.B. and Zinder, N.D. (1972), *Virology*, 49, 45-60.
35. Sugiura, M., Okamoto, T. and Takanami, M. (1969), *J. Mol. Biol.*, 43, 299-315.

CHAPTER V

A CASCADE MECHANISM OF TRANSCRIPTION IN BACTERIOPHAGE M13 DNA

A CASCADE MECHANISM OF TRANSCRIPTION IN BACTERIOPHAGE M13 DNA

ABSTRACT

Transcription studies were performed on M13 DNA restriction fragments which encompass a single rho-independent termination site and one or more promoter sites. From the results of these studies it has been concluded that a cascade-like mechanism of transcription operates during the expression of the phage M13 genome. This implies that each promoter gives rise to the formation of an RNA chain whose synthesis is not terminated before the central termination site for transcription is reached. Most genes are therefore read into primary transcripts that originate from several different promoters.

By means of uncoupled, as well as coupled transcription and translation studies, it has been demonstrated that the RNA species initiated at the five, previously identified G-promoters (i.e. $G_{0.82}$, $G_{0.88}$, $G_{0.94}$, $G_{0.01}$ and $G_{0.08}$) are, respectively 360, 750, 1200, 1600 and 2000 nucleotides long. Furthermore, it has been shown that the promoters $G_{0.82}$, $G_{0.88}$ and $G_{0.08}$ are located immediately in front of gene VIII, gene V and gene II, respectively, and that the other two G-promoters ($G_{0.94}$ and $G_{0.01}$) are intragenic RNA initiation sites which are located within gene II. In addition, it has been demonstrated that promoter $G_{0.94}$ is located immediately in front of the DNA region which codes for the synthesis of X-protein.

Electrophoretic analysis of the products transcribed from fragments which contained the central termination signal have indicated that the termination of transcription is not stringent. The implication of this read-through phenomenon with respect to the expression of the genes located immediately distal to the central termination is discussed.

INTRODUCTION

The small filamentous coliphages M13, fd and f1 provide useful models for the study of a number of mechanisms by which genes are expressed. Their genome consists of a circular single-stranded DNA molecule (mol. wt. ca. 2×10^6) on which at least eight genes are located (for recent reviews, see Denhardt, 1975 and Ray, 1976). The physical order of these genes has been described recently (van den Hondel *et al.*, 1975; Vovis, Horiuchi and Zinder, 1975). After infection, the single-stranded phage DNA is converted into a circularly closed double-stranded replicative form molecule (RF-I). Both *in vivo* as well as *in vitro*,

only the nonviral strand of this duplex DNA molecule functions as a template for transcription (Okamoto, Sugiura and Takanami, 1969; Edens et al., 1976; Jacob, Jaenisch and Hofschneider, 1970; Konings, Smits and Schoenmakers, submitted for publication.

Evidence has accumulated that the M13 genome is expressed in vitro as well as in vivo according to a cascade-like mechanism of transcription (Edens, Konings and Schoenmakers, 1975; Chan, Model and Zinder, 1975; Okamoto et al., 1975; Edens et al., 1976; Konings, Smits and Schoenmakers, submitted for publication). Such evidence implies that RF DNA is transcribed into a discrete number of RNA species which are initiated at different promoter sites (RNA initiation sites) but which are all terminated at a unique termination signal. A direct consequence of this mechanism of gene regulation is that genes which are located proximal to the 5'-end of this termination signal are transcribed more frequently than those which are located at greater distances.

Recently, it has been reported that a strong, rho-independent termination signal is located immediately after the 3'-end of gene VIII (Edens, Konings and Schoenmakers, 1975; Sugimoto et al., 1976). Furthermore, we have mapped eight promoter sites on the M13 genome (Edens et al., 1976). Five of these promoters, designated $G_{0.82}$, $G_{0.88}$, $G_{0.94}$, $G_{0.01}$ and $G_{0.08}$, which produce pppG-initiated RNA chains, were found to be clustered proximal to the 5'-end of the termination signal, in a region which encompasses only one third of the genome (Edens, Konings and Schoenmakers, 1975; Edens et al., 1976). The other promoters, all of which produce pppA-initiated RNA chains, were found to be localized at greater distances from this termination signal.

With respect to the regulation of the expression of the M13 genome, we have studied the validity of the cascade-transcription model in more detail. For this reason we have searched for restriction fragments of M13 RF which contain both the rho-independent termination signal as well as one or more G-promoters. Subsequently, the properties of these fragments were investigated with respect to their capability of directing the synthesis of RNA in vitro. Furthermore, it was determined whether or not these fragments were able to direct the synthesis of phage-spe-

cific polypeptides in a coupled transcription and translation system. It has been demonstrated that the region of the M13 genome under study is in fact under the control of a cascade-like mechanism of transcription. These studies have enabled us, furthermore, to locate several promoter sites more precisely.

MATERIALS AND METHODS

Enzymes

Escherichia coli RNA polymerase holoenzyme was a generous gift from Dr. R. Schilperoort (University of Leiden). The restriction endonucleases from Haemophilus influenzae Rd (endo R. Hind II), Haemophilus aphirophilus (endo R. Hap II), Haemophilus aegyptius (endo R. Hae II) and Haemophilus influenzae Rf (endo R. Hinf I) have been isolated as described previously (van den Hondel and Schoenmakers, 1973, 1975 and 1976; van den Hondel, Pennings and Schoenmakers, 1976).

Replicative form I DNA and restriction fragments

The method for the preparative isolation and purification of M13 replicative form I DNA has been described (van den Hondel and Schoenmakers, 1976). Digestion of it with the restriction endonucleases has been described elsewhere (van den Hondel and Schoenmakers, 1975). After digestion, the fragments were separated by electrophoresis on 2.5% (for endo R.Hae II fragments) or 3.0% discontinuous polyacrylamide slab gels and further purified as described by van den Hondel et al. (1975).

RNA synthesis in vitro

RNA synthesis in vitro was performed in a standard reaction mixture (0.1 ml) which contained 4 µmol Tris-HCl (pH 7.9), 15 µmol KCl, 0.8 µmol $MgCl_2$, 0.1 µmol dithitreitol, 0.01 µmol EDTA, 0.1 µmol UMP, 0.1% Tween-80, 25 µg bovine serum albumin, 0.2 pmol M13 DNA restriction fragment and 8.0 pmol E.coli RNA polymerase holoenzyme. After a pre-incubation period of 5 min at 37°C, ribonucleoside triphosphates were added to a final

concentration of 80 μM, except for the (α-^{32}P)-labeled CTP whose concentration was 8 μM. After further incubation for 10 min the reactions were terminated by the addition of 10 μl 10% SDS and 0.1 ml phenol saturated with a buffer containing 0.01 M Tris-HCl (pH 7.6) and 0.001 M EDTA. M13 RF was similarly transcribed with the exceptions that the molar ratio of RNA polymerase to DNA template was 10, and a five fold excess of ribonucleoside triphosphates was added. After RNA synthesis, 10 μg of carrier tRNA was added and the reaction mixtures were extracted with phenol. Subsequently the RNA was precipitated twice with 2.5 volumes of cold ethanol and finally dissolved in 15 μl of deionized formamide.

Polyacrylamide gel electrophoresis

The in vitro synthesized RNA species were analysed on 1.9% polyacrylamide slab gels in the presence of 7 M urea (Konings and Bloemendal, 1969). After electrophoresis for 5 h at 50 mA, the wet gel slab was exposed to an X-ray film (Kodak RP/R54). The approximate lengths of the RNA species synthesized in vitro were estimated from their relative electrophoretic mobilities. E.coli ribosomal RNAs and denatured restriction fragments of M13 RF were used as electrophoretic markers. The chain lengths for the various RNAs were assumed to be 3000 nucleotides for 23S, 1500 nucleotides for 16S and 120 nucleotides for 5S ribosomal RNA (Fellner, 1974). For the restriction fragment lengths the data of van den Hondel and Schoenmakers (1975) were consulted. No detectable differences were observed between the relative electrophoretic mobilities of the ribosomal RNAs and the restriction fragments of identical length.

Isolation of RNA species

Radioactive bands were cut from the frozen gel and twice extracted at room temperature for two hours in a buffer containing 0.1 M NaCl, 0.01 M Tris-HCl (pH 7.6), 0.001 M EDTA and 0.1% SDS. After low speed centrifugation the supernatant was collected and the RNA was precipitated by the addition of 0.1

volume of 3 M sodiumacetate (pH 5.3) and 2.5 volumes of cold ethanol. For *in vitro* protein synthesis studies, the RNA was purified by repeated centrifugation in glycerol gradients (5-30% glycerol in a buffer containing 0.1 M NaCl, 0.01 M Tris-HCl (pH 7.6), 0.001 M EDTA and 0.1% SDS). Following this purification procedure the recovering of the RNA was approximately 70%.

Protein synthesis

Cell-free protein synthesis, under the direction of restriction fragments or under the direction of the *in vitro* synthesized mRNA species, was accomplished as described previously (Konings, Ward, Francke and Hofschneider, 1970; Konings, Hulsebos and van den Hondel, 1975). The synthesized polypeptides were analysed on 15% SDS-Tris-glycine polyacrylamide gels in the presence of 8 M urea (Konings, Hulsebos and van den Hondel, 1975).

Nomenclature of promoter sites

Promoters which initiate the synthesis of RNA chains beginning with pppG are denoted by G, while those which initiate the synthesis of RNA beginning with pppA are denoted by A. The position of each promoter is given in map units and is indicated by a suffix which corresponds to the position of the promoter on the physical map.

RESULTS

In previous communications we have reported the construction of several enzyme cleavage maps of M13 replicative form DNA and the correlation of each map with the genetic map of the M13 genome (van den Hondel *et al.*, 1975; van den Hondel, Pennings and Schoenmakers, 1976). Also the map positions of eight promoter sites and a strong, rho-independent termination signal have been described (van den Hondel, Konings and Schoenmakers, 1975; Edens, Konings and Schoenmakers, 1975; Edens *et al.*, 1976). The positions of these regulatory elements on the physical map are

indicated in Fig. 1. From this figure it furthermore can be deduced that a number of restriction fragments are available which contain, in addition to the rho-independent termination signal, one or more G-promoters. This has enabled us to study in more detail whether a cascade mechanism of transcription operates during the expression of this particular part of the M13 genome. For this reason each of these fragments was tested as a template for transcription by E.coli RNA polymerase holoenzyme. After incubation the synthesized RNA chains were fractionated by means of polyacrylamide gel electrophoresis and their lengths were estimated. Subsequently, they were extracted from the gel and after further purification (Materials and Methods),were tested for their capacity to direct the in vitro synthesis of phage-specific polypeptides. To verify the results obtained from the uncoupled transcription and translation studies, the DNA fragments were also directly used as a template in a coupled transcription and translation system; the results of these studies are discussed below.

Fragment Hap II-B_2

Restriction fragment Hap II-B_2 is approximately 800 base pairs in length (Van den Hondel and Schoenmakers, 1975). Genetic analysis and in vitro transcription and translation studies have indicated that this fragment encompasses genes VII and VIII and the N-terminal region of gene III (Van den Hondel et al., 1975; Van den Hondel, Konings and Schoenmakers, 1975). Previously, we have demonstrated that both promoter $G_{0.82}$ as well as the rho-independent termination signal for transcription are located on this restriction fragment (Edens, Konings and Schoenmakers, 1975). These regulatory elements were found to be located at approximately 90 and 450 nucleotides, respectively, form the 5'-end of this restriction fragment † (Fig. 1).

In accordance with these observations we have found that fragment Hap II-B_2 directs in vitro the synthesis of an RNA species which is about 360 nucleotides in length (8S RNA; Fig. 2a). Since fragment Hap II-B_2 also directs the synthesis of the precursor of the capsid protein (mol. wt. 7500) encoded by gene VIII (Van den Hondel, Konings and Schoenmakers, 1975; Edens, Konings and

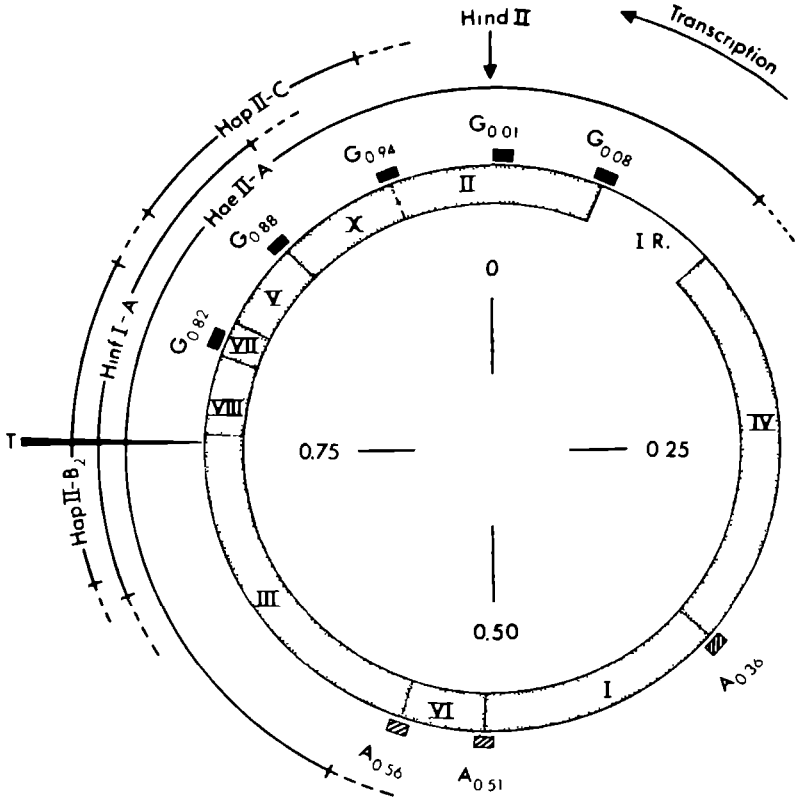


Fig. 1: Genetic map and physical maps of bacteriophage M13 DNA. The dotted circular area represents the genetic map. The uncompleted outer circles show the locations of restriction fragments which cover the central termination site of transcription (T) and one or more promoter sites (G). The position of the promoters discussed in this paper are indicated: G-promoters with black bars, and A-promoters with hatched bars. The direction of transcription is counterclockwise around the genetic map. The arrow indicates the cleavage site of M13 RF by endo R. Hind II. 'X' refers to the C-terminal part of gene II in which an internal start for RNA synthesis is located. This RNA directs the synthesis of protein X. I.R. refers to the intergenic region between gene II and gene IV in which the replication origin for parental RF is located.

† Note: The direction of transcription proceeds counterclockwise on the genetic map. This is also the 5' to 3' polarity direction in the viral strand of each restriction fragment.

Schoenmakers, 1975; cf. Fig. 3a), this transcription product should also direct the synthesis of the latter polypeptide. As shown in Fig. 4f, this indeed could be demonstrated. From these observations we conclude, therefore, that promoter $G_{0.82}$ is located immediately in front of gene VIII.

As shown in Fig. 2a, fragment Hap II-B_2 also directs, besides the synthesis of 8S RNA, the weak synthesis of an RNA species which is about 720 nucleotides (ca. 11S) in length. The length of the latter RNA product corresponds to the distance between promoter $G_{0.82}$ and the 3'-end of fragment Hap II-B_2 (Fig. 1). This fact, together with the observation that it directs the synthesis of gene VIII protein (data not shown) suggests that this 720 nucleotides-long RNA is the result of a transcription process which is initiated at promoter $G_{0.82}$ but which is not terminated until the RNA polymerase molecule has reached the 3'-end of this restriction fragment. From these results we conclude therefore, that the termination of transcription at the rho-independent termination signal is not stringent but that the RNA polymerase molecule frequently 'leaks' through this termination site. From the ratios of the relative amounts of both synthesized RNA products we have calculated that about ten percent of the RNA polymerase molecules read through the rho-independent termination signal.

Fragment Hinf I-A

Fragment Hinf-A is about 1320 base pairs in length (Van den Hondel and Schoenmakers, 1976). As judged from its position on the physical map this fragment encompasses genes VIII, VII and V and the regions coding for the N-terminal end of gene III and the C-terminal end of gene II (Fig. 1). Hence, this fragment is presumed to contain not only the rho-independent termination signal but also the promoters $G_{0.82}$ and $G_{0.88}$.

If in fact a cascade mechanism of transcription operates during the expression of the M13 genome, then this fragment should direct the synthesis of at least two RNA species of different length. In addition these RNA species should have in common the coding information for the polypeptide encoded by gene VIII. As shown in Fig. 2b, upon transcription, fragment Hinf I-A gives

rise to the formation of three different RNA chains of which the smallest co-migrates with the 8S RNA transcribed either from fragment Hap II-B_2 (Fig. 2a) or from replicative form DNA (Fig. 2e). The two larger RNA species have electrophoretic mobilities which correspond to chain lengths of about 750 and 800 nucleotides. Only the 750 nucleotide transcript has an equivalent among the RNA species (i.e. 11S RNA) transcribed from the intact M13 genome (Fig. 2e). From these data, together with the results described previously, we conclude that the 8S and the 11S RNA are initiated at the promoters $G_{0.82}$ and $G_{0.88}$, respectively.

The length of the 800 nucleotide transcript corresponds to the distance between promoter $G_{0.82}$ and the 3'-end of fragment Hinf I-A (Fig. 1). Similar to what has been found in the case of the 720 nucleotide transcript of fragment Hap II-B_2 (previous section), this observation also suggests that termination of transcription at the rho-independent termination signal is not stringent. Although one also should expect to find a read-through product of the RNAs initiated at promoter $G_{0.88}$, this has never been found (cf. Fig. 2b). A possible reason for its apparent absence is that the affinity of RNA polymerase holoenzyme for promoter $G_{0.88}$ is at least five times lower than for promoter $G_{0.82}$ (Edens et al., 1976). As a consequence of this, small amounts of read-through products are synthesized and therefore may escape detection or the gel systems used.

To substantiate the transcription data, fragment Hinf I-A was also used as a template in a coupled in vitro protein synthesizing system. As shown in Fig. 3b, this fragment directs the synthesis of the polypeptides encoded both by genes VIII and V. A similar result was obtained when the purified 8S RNA and 11S RNA species, transcribed in vitro from fragment Hinf I-A, were translated (Figs. 4e and 4f). The smallest RNA (8S) codes for gene VIII protein while the larger RNA (11S) codes for both gene VIII- and gene V-protein (Figs. 4e and 5b). From these results we conclude that apart from the promoter in front of gene VIII ($G_{0.82}$) a second promoter ($G_{0.88}$) is present on fragment Hinf I-A which is located immediately in front of gene V.

Although the complete gene VII is located on fragment Hinf I-A (Fig. 1), for reasons still unknown, we have been unable to

Fig. 2: Autoradiogram of ^{32}P-labeled RNA species synthesized *in vitro* under the direction of either phage M13 replicative form DNA or various restriction fragments of M13 RF. Products transcribed from (a) fragment *Hap* II-B_2; (b) fragment *Hinf* I-A; (c) the mixture of the fragments *Hin*. *Hae* II-A_1 and *Hin*. *Hae* II-A_2; (d) fragment *Hae* II-A; (e) wild-type replicative-form DNA; (f) fragment *Hap* II-C. RNA synthesis was carried out in the presence of (α-^{32}P)-labeled CTP and was performed as described under Materials and Methods. The RNA products were analyzed on vertical slab gels (1.9% acrylamide) in the presence of 7M urea. The sizes of the RNA products made were estimated as described previously (Edens *et al.*, 1976).

demonstrate the in vitro synthesis of its encoded polypeptide both in a coupled and in an uncoupled cell-free protein synthesizing system (Van den Hondel, Konings and Schoenmakers, 1975; Edens, Konings and Schoenmakers, 1975). In fact, at present, and most likely for identical reasons, we have not been able to demonstrate, both in vivo as well as in vitro, the synthesis of gene VII protein either under the direction of the intact M13 genome or under the direction of its transcripts (Konings, Hulsebos and Van den Hondel, 1975; Model and Zinder, 1974; Smits and Schoenmakers, submitted for publication). Since both the coupled and the uncoupled protein synthesis studies were carried out at a ribosome excess, the absence of gene VII-protein synthesis can not be attributed to a shortage of ribosomes needed for the translation of the gene VII cistron.

Fragment Hap II-C

As shown in Fig. 1, fragment Hap II-C which is about 650 base pairs in length (Van den Hondel and Schoenmakers, 1975), partially overlaps the 5'-end of fragment Hinf I-A. Previously, we have demonstrated that this fragment directs, in a DNA-dependent protein synthesizing system, the synthesis of a polypeptide with a molecular weight of 12000 (Van den Hondel, Konings and Schoenmakers, 1975). Although M13-specified, this polypeptide could not be attributed to one of the known M13 genes, and was therefore arbitrarily called "X-protein". Recent results, however, have indicated that X-protein is the product of a gene, i.e. gene X, which is located completely within another gene, namely gene II and that the nucleotide sequence of that particular gene overlaps with the C-terminal end of gene II (Van den Hondel, Pennings and Schoenmakers, 1976). From transcription and translation data it could be deduced further that immediately in front of gene X a promoter site ($G_{0.94}$) is located (Van den Hondel, Konings and Schoenmakers, 1975; Edens et al., 1975).

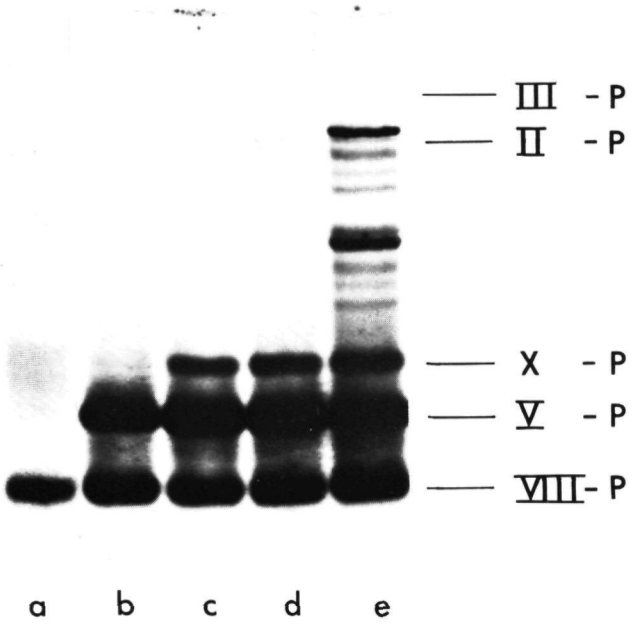


Fig. 3: Autoradiogram of (^{35}S) methionine-labeled polypeptides synthesized *in vitro* under the direction of various restriction fragments of phage M13 replicative form DNA. Polypeptides synthesized under the direction of (a) fragment *Hap* II-B_2; (b) fragment *Hinf* I-A; (c) the mixture of the fragments *Hin*. *Hae* II-A_1 and *Hin*. *Hae* II-A_2; (d) fragment *Hae* II-A; (e) wild-type M13 replicative-form DNA. Protein synthesis *in vitro* was carried out as described (Konings *et al.*, 1975). Of each DNA fragment, an amount equivalent to 1 µg of replicative-form DNA was added per 0.025 ml of reaction mixture. The *in vitro* synthesized polypeptides were analyzed on 15% SDS-Tris-glycine gels (Konings *et al.*, 1975). The identification and characterization of the products synthesized *in vitro* under the direction of wild-type replicative-form DNA have been described (Konings *et al.*, 1975). I-P, II-P, etc. refer, respectively to the positions of migration of the products of gene I, gene II, etc.

Transcription of fragment Hap II-C results in the formation of a major RNA species of about 570 nucleotides, and two minor species which are about 180 and 1300 nucleotides in length (Fig. 2f). From the distance between promoter $G_{0.88}$ and the rho-independent termination signal (previous section) and from the lengths of the various restriction fragments, it can be deduced that promoter $G_{0.88}$ is located at a distance of about 160 nucleotides from the 3'-end of fragment Hap II-C. The formation of an RNA transcript of Hap II-C of approximately this size accords with such a promoter position. Since fragment Hap II-C also directs the synthesis of an RNA species which is about 570 nucleotides in length we suggest that yet another promoter ($G_{0.94}$) is located on fragment Hap II-C. The presence of a G-promoter at this position has already been demonstrated unambiguously by Takanami and co-workers (Sugimoto et al., 1975) and by our group (Van den Hondel, Konings and Schoenmakers, 1975; Edens, Konings and Schoenmakers, 1976). This promoter, in fact, is the strongest promoter located on the M13 genome. Its high affinity for RNA polymerase holoenzyme can be recognized readily by comparison of the differences in intensities between the various bands present on the autoradiographs of the RNAs synthesized under direction of fragment Hap II-C (Fig. 2f). The same is true for the 14S RNA, i.e. the RNA product which is initiated at promoter $G_{0.94}$ and transcribed from the intact M13 genome (Fig. 2e).

Previously, we have reported (Edens et al., 1976) that in the absence of a natural termination signal on a restriction fragment, termination of transcription occurs at the 3'-end. According to Takanami's data (personal communication), the RNA polymerase molecule at this end most probably switches transcription from one strand to the other and forms a small hairpin loop. This, in fact, would be the signal for the RNA polymerase molecule to stop transcription. Infrequently, however, minor transcripts are formed on this fragment which are one (or more) fragment length (s) greater than the major transcripts. A typical example is the 1300 nucleotide-long transcript (Fig. 2) which is about 700 nucleotides longer than the major transcript (570 nucleotides) initiated at promoter $G_{0.94}$ on fragment Hap II-C. How these extended RNA species are formed is not known.

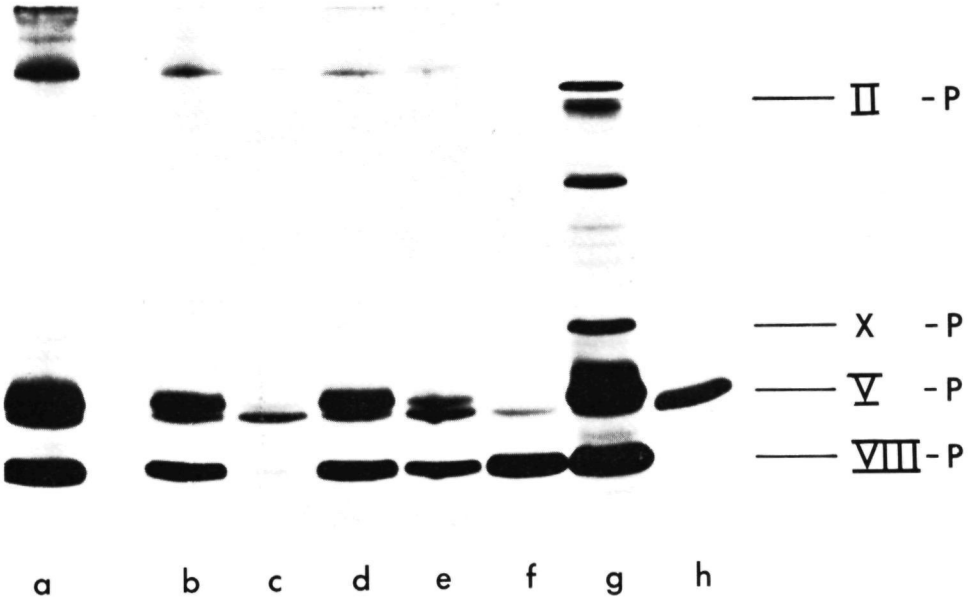


Fig. 4: Autoradiogram of (^{35}S) methionine-labeled polypeptides synthesized in vitro under the direction of the individually purified mRNA species transcribed from various restriction fragments of phage M13 DNA. Products synthesized under the direction of the (a) 14S; (b) 19S; (c) 17S; (d) 14S; (e) 11S and (f) 8S RNA species; (g) products synthesized under the direction of wild-type phage replicative form DNA; (h) products synthesized in the absence of exogeneous RNA or DNA. RNA synthesis was carried out as described under Materials and Methods. The in vitro-synthesized RNA molecules were fractionated by means of electrophoresis on 1.9% polyacrylamide gels in the presence of 7 M urea. After radioautography, the radioactive bands were cut from the frozen gel and after elution and additional purification the different RNA classes were translated into protein as described by Konings et al., (1975). In case protein synthesis was studied under the direction of the in vitro-synthesized RNA species, 100 µg of rifampycin/ml was added. The in vitro - synthesized polypeptides were analyzed on 15% SDS-Tris-glycine gels (Konings et al., 1975). The identification and characterization of the products synthesized in vitro under the direction of wild-type replicative form DNA have been described (Konings et al., 1975). I-P, II-p, etc. refer, respectively, to the positions of migration of the products of gene I, gene II, etc. In case of lane a the film was exposed for a longer period of time than in the other situations.

One explanation would be that, infrequently, cessation of transcription at the 3'-end of this fragment does not occur but after the loop has formed transcription proceeds along the fragment in the opposite direction.

As is evident from previous communications (Van den Hondel, Konings and Schoenmakers, 1975; Van den Hondel, Pennings and Schoenmakers, 1976), the genetic information which codes for the synthesis of X-protein is located between promoters $G_{0.94}$ and $G_{0.88}$ (Fig. 1). This polypeptide is in fact the sole polypeptide synthesized under the direction of fragment Hap II-C when it is added as a template to the DNA-dependent protein synthesizing system (Van den Hondel, Konings and Schoenmakers, 1975; cf. Fig. 5f).

Fragment Hin. Hae II-A

Endonuclease R. Hind II has only one cleavage site on M13 RF (Van den Hondel and Schoenmakers, 1973 and 1975). This site is located within gene II at a position which has been defined as the zero point of the physical map (Fig. 1). With the aid of this enzyme and endonuclease R. Hae II two fragments of M13 RF can be produced with one Hae II cleavage site at one end and a Hind II cut at the other. These fragments, designated Hin.Hae II-A_1 and Hin. Hae II-A_2, are about 2700 and 800 base pairs in length (Van den Hondel, Pennings and Schoenmakers, 1976). Previously, we have reported that fragment Hin. Hae II-A_1 covers genes VIII, VII and V, the N-terminal part of gene III and also a considerable part of gene II (Fig. 1; Van den Hondel *et al.*, 1975; Van den Hondel, Pennings and Schoenmakers, 1976). The N-terminal part of gene II is not contained in this fragment but is present in the small Hin. Hae II-A_2 fragment. Therefore, it is presumed that the large fragment contains the rho-independent termination signal and the promoters $G_{0.82}$, $G_{0.88}$ and $G_{0.94}$, whereas the small fragment only contains promoters $G_{0.01}$ and $G_{0.08}$ (Fig. 1).

Electrophoretic analysis of the RNA products formed by transcription of the unfractionated fragments Hin. Hae II-A_1 and Hin. Hae II-A_2 has indicated that they direct the synthesis of six different RNA species (Fig. 2c). The major products are 8S, 9S,

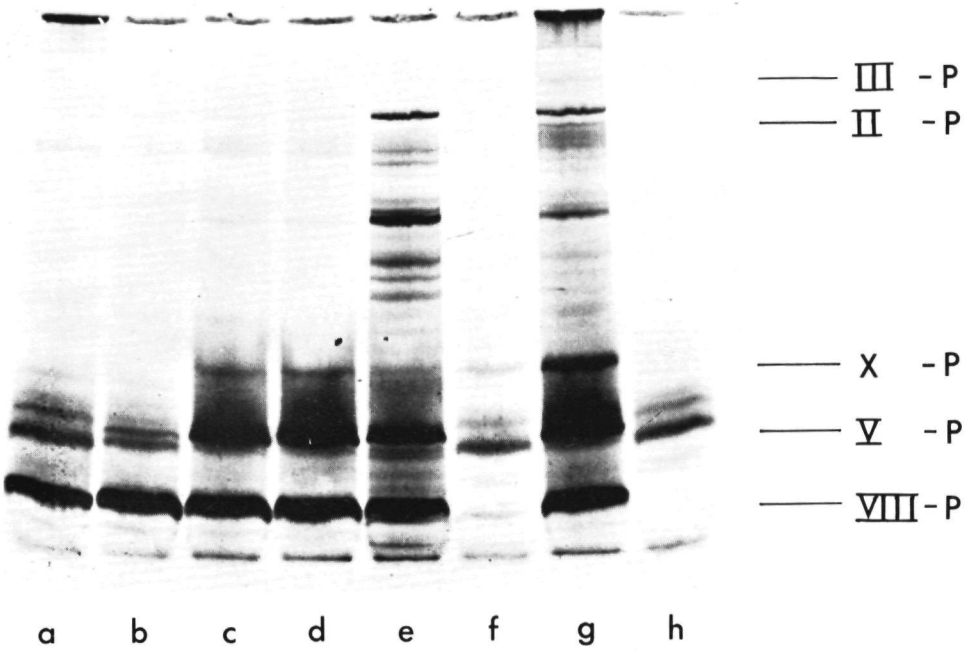


Fig. 5: Autoradiogram of (^{35}S) methionine-labeled polypeptides synthesized *in vitro* under the direction of the mRNA species transcribed from various restriction fragments of phage M13 replicative-form DNA. Products synthesized under the direction of RNA species transcribed (a) from restriction fragment *Hap* II-B_2; (b) from fragment *Hinf* I-A; (c) from the mixture of restriction fragments *Hin*. *Hae* II-A_1 and *Hin*. *Hae* II-A_2; (d) from fragment *Hae* II-A; (e) from wild-type M13 replicative form DNA; (f) from fragment *Hap* II-C; (g) products synthesized under the direction of wild-type M13 replicative form DNA; (h) products synthesized in the absence of exogeneous RNA or DNA.

RNA synthesis was carried out as described under Materials and Methods. After purification the *in vitro*- synthesized RNA species were translated into protein as described in the legend to Fig. 4. The identification and characterization of the products synthesized *in vitro* under the direction of wild-type replicative form DNA have been described (Konings *et al.*, 1975). I-P, II-P, etc. refer, respectively, to the positions of migration of the products of gene I, gene II, etc.

11S and 14S RNA. In addition, two minor RNA products are formed which are larger than the 14S RNA. The lengths of these two RNA chains correspond to the distance between the strong promoters $G_{0.82}$ and $G_{0.94}$ on one side and the 3'-end of fragment Hin. Hae II-A_1 on the other. As discussed in the previous sections, these RNA species, therefore, are most probably the result of 'leakage' of the RNA polymerase molecule through the rho-independent termination signal.

Considering the results described in the previous paragraphs, it seems justified to conclude that the 8S, 11S and 14S RNA species are initiated at the promoters $G_{0.82}$, $G_{0.88}$ and $G_{0.94}$, respectively, and that they are all terminated at the same termination signal located immediately after gene VIII. The 9S RNA species most probably represents a product which is initiated at promoter $G_{0.08}$ and which is prematurely terminated at the 3'-end of the small fragment Hin. Hae II-A_2 (cf. Edens et al., 1976).

From the data presented so far, it is clear, that within the region of the M13 genome under study, natural termination of transcription does not occur at positions other than the previously identified, rho-independent termination signal (Edens, Konings and Schoenmakers, 1975), which is located immediately after gene VIII. Most RNA species are thus the result of transcription of RNA polymerase holoenzyme through promoter sites which are located between their own RNA initiation site and the central termination site of transcription. The results obtained from coupled and uncoupled in vitro protein synthesis studies are in accordance with such a reading of the RNA polymerase molecule through promoter sites, ultimately leading to cascade transcription. As shown in Fig. 3c, the mixture of the fragments Hin. Hae II-A_1 and Hin. Hae II-A_2 direct the synthesis of X-protein as well as the polypeptides encoded by genes V and VIII. In Fig. 5c a pattern is shown of the in vitro products produced when the mixture of RNA species transcribed from these fragments was added to the uncoupled protein synthesizing system. This pattern is almost identical to that found when the DNA fragments were added directly to the coupled cell-free system. An exception, however, is X-protein. This polypeptide is a prominent product in the DNA-dependent system, whereas it is made in much reduced amounts when its synthesis

is studied under uncoupled conditions (cf. Figs. 3c and 5c). The reason why the synthesis of X-protein is drastically reduced under uncoupled conditions is not known.

From the results presented in the previous sections, it is clear that the 8S RNA and the 11S RNA species share the coding information for the polypeptide encoded by gene VIII. If a cascade mechanism of transcription is operative during the expression of the M13 genome, then the 14S RNA species should also direct, besides the synthesis of X-protein and gene V protein, the synthesis of this polypeptide. As shown in Fig. 4a (cf. Fig. 4d) this proved to be the case. Similar to what has been found in the case of translation of total RNA transcribed from RF-I DNA (Fig. 5e) also in this situation X-protein is synthesized in relatively smaller amounts than when it is made in the DNA-dependent protein synthesizing system under the direction of restriction fragments or under the direction of replicative form DNA (Fig. 3e; cf. Van den Hondel, Konings and Schoenmakers, 1976).

Hae II-A

Endonuclease R. Hae II cleaves M13 DNA into three fragments, of which the largest fragment Hae II-A is slightly larger than half genome length (Horiuchi et al., 1975; Van den Hondel, Pennings and Schoenmakers, 1976). It encompasses a very large part of gene III, the genes II, V, VII and VIII entirely and a small part of the C-terminal end of gene IV. In addition, the intergenic region (IR) which encompasses the origin of M13 RF replication, is located on this fragment (Horiuchi and Zinder, 1976; Van den Hondel, Konings and Schoenmakers, submitted for publication; Suggs and Ray, 1977). As shown in Fig. 1, this region is located between the C-terminal part of gene IV and the N-terminal part of gene II.

Transcription of fragment Hae II-A results in the formation of a discrete number of RNA species of different lengths (Fig. 2d). The five major products, namely 8S, 11S, 14S, 17S and 19S RNA co-migrate exactly whith mRNA species transcribed from intact replicative form DNA (Fig. 2e). These observations, together with the results described in the previous sections, suggest

that the smaller RNA species, 8S, 11S and 14S RNA, are the products which are initiated at the promoters $G_{0.82}$, $G_{0.88}$ and $G_{0.94}$, respectively and that the largest RNA species (i.e. 19S RNA) is initiated at the promoter $G_{0.08}$ which is most probably located immediately in front of gene II. Our previous observation of a second RNA initiation site within gene II (i.e. $G_{0.01}$), is now further supported by the observed synthesis of a 17S RNA species under the direction of both M13 RF and fragment Hae II-A (Figs. 2d and 2e). This intragenic start of RNA synthesis is, however, not observed when fragment Hae II-A is cleaved once with the Hind II enconuclease (Fig. 2c). This suggests that the promoter $G_{0.01}$ is located so close to the Hind-II cut that after cleavage its capacity to initiate RNA synthesis is destroyed completely.

The validity of our conclusions is further substantiated by coupled and uncoupled *in vitro* protein synthesis studies. As shown in Fig. 3d, fragment Hae II-A directs, in the DNA-dependent protein synthesizing system, not only the synthesis of X-protein and the polypeptides encoded by genes V and VIII, but also the synthesis of gene II protein. Hence our previous assumption (Edens *et al.*, 1976) that a promoter site ($G_{0.08}$) is located immediately in front of gene II is proved to be valid.

Similar results were obtained when the unfractionated RNA species transcribed from fragment Hae II-A were used as a messenger to direct protein synthesis in the uncoupled cell-free system (Fig. 5d). Again a prominent synthesis of gene VIII protein and gene V protein is apparent, whereas, in comparison to the products made in the coupled system (Fig. 3d), the synthesis of X-protein and gene II protein is reduced greatly. Up to now, we have not been able to demonstrate the synthesis of a polypeptide which is encoded by the region located between the promoters $G_{0.01}$ (17S) and $G_{0.94}$ (14S) (cf. Figs. 3c, 3d, 4c and 4d). The reason why upon transcription of M13 RF two different RNA chains are made, whose coding information is, at least in our hands, nevertheless identical, cannot be answered yet.

On the basis of the results of *in vitro* transcription studies, it has been suggested by Okamoto, Sugiura and Takanami (1969) that the replicative form DNA of the filamentous coliphages may contain several promoter sites but only one termination signal. Since then, a number of studies from several laboratories (Seeburg and Schaller, 1975; Chan, Model and Zinder, 1975; Okamoto *et al.*, 1975) as well as from ours (Edens, Konings and Schoenmakers, 1975; Edens *et al.*, 1976; Konings, Edens and Schoenmakers, submitted for publication) have supported and further extended this transcription model.

To investigate this transcription mechanism in more detail in this study, we have made use of restriction fragments of M13 RF. The restriction fragments were so selected that they contained, besides the previously identified rho-independent termination signal (Edens, Konings and Schoenmakers, 1975), one or more promoter sites. The RNA products synthesized by RNA polymerase holoenzyme were fractionated on polyacrylamide gels and after extraction and further purification each RNA species was used as a template to direct the *in vitro* synthesis of phage M13 specific proteins. The results obtained are completely consistent with a multiple-promoter, single-terminator model. Most of the RNA chains initiated at a particular promoter site are thus the result of transcription through other promoters until, finally, the central termination signal for transcription is reached.

Apart from the observed differences in transcriptional activity and strength of promoter selection (Edens *et al.*, 1976; Seeburg, Nüsslein and Schaller, 1976), cascade-transcription on a circular genome leads to a stepwise gradient of RNA molecules with overlapping nucleotide sequences, with maximal overlap immediately proximal to the central terminator. Our previous reports that genes V and VIII are positioned in the immediate proximity of this terminator (Edens, Konings and Schoenmakers, 1975; Van den Hondel *et al.*, 1975), now provides a qualitative explanation for the observed abundant *in vitro* synthesis of gene V and of gene VIII protein (cf. Konings, 1973; Model and Zinder, 1974; Konings, Hulsebos and Van den Hondel, 1975). In this connection

it should be mentioned that the latter two proteins are also the major M13 products made in the infected cell (Henry and Pratt, 1969; Konings, Smits and Schoenmakers, submitted for publication). Furthermore, it has been observed that most of the *in vitro* synthesized G-start RNA's have an equivalent among the *in vivo* synthesized mRNA species (Konings, Smits and Schoenmakers, submitted for publication). For these reasons it is reasonable to assume that the mechanism of cascade transcription is also operative *in vivo*. If this assumption is correct, then the observation that during the phage infection cycle only small amounts of gene III protein are made can be explained, at least in part, by the fact that its corresponding gene is positioned on the genetic map immediately distal to the central terminator (Fig. 1).

The overall synthesis of phage M13 specific proteins, however, is not solely regulated at the level of transcription. One might expect that, once the primary transcripts have been formed, the rate of synthesis is also influenced by additional factors, like differences in binding of ribosomal subunits at the protein synthesis initiation sites, the accessibility of such sites and the biological half-lives of such messages. That such circumstantial factors must be taken into account can be deduced from the finding that the purified individual mRNA species direct *in vitro* the synthesis of larger amounts of gene VIII protein (and of gene V protein) than the other phage M13 encoded polypeptides (Fig. 4; cf. Fig. 5).

Furthermore, hitherto, we have been unable to demonstrate the *in vitro* synthesis of gene VII protein although the coding information for this polypeptide is present on all RNA chains which are larger than the 8S RNA species. Regulation at the level of translation is also apparent during the expression of the genes coding for X protein and for gene II protein. The synthesis of the latter two proteins is less pronounced if protein synthesis is directed by RNA instead of DNA (cf. Figs. 3 and 4). The reason for this is not clear. It is possible that the ribosomes have a higher affinity for the ribosomal binding sites present on the nascent mRNA chains than for those present on the completed polycistronic mRNA molecules.

Recently, we have demonstrated the existence of five G-promoters and three A-promoters on replicative form M13 DNA (Konings, Hulsebos and Van den Hondel, 1975; Edens et al., 1976). The positions of the five G-promoters have been confirmed in this study and are indicated in Fig. 1. Seeburg and Schaller (1975), Okamoto et al., (1975) and Chan, Model and Zinder (1975) have studied promoter locations on the genomes of the phages M13, fd and f1. There exists excellent agreement between the results of our laboratory and those of Seeburg and Schaller, except for the presence of the $G_{0.08}$ promoter and the $A_{0.51}$ promoter, which is most probably located in front of gene I. As discussed previously (Edens et al., 1976), these differences in results most probably reflect differences in techniques used for the mapping of promoter sites. The results of Okamoto et al. (1975) are only consistent with ours as far as the $G_{0.82}$, $G_{0.94}$ and $G_{0.08}$ promoters are concerned. Agreement is poor with respect to the A-promoters, whereas the existence of the promoters $G_{0.88}$ and $G_{0.01}$ has not been observed by this group. Both on the basis of the lengths of the RNA chains transcribed from f1 RF as well as on the basis of in vitro protein synthesis studies, Chan, Model and Zinder (1975) have made estimates of the location of several f1 promoters. Unequivocal data, based on transcription studies on DNA fragments which are presumed to contain these promoters, have not been given, however. Hence, the promoter positions reported by this group have to be considered as rough approximations.

There exists, however, general agreement concerning the positions of the strong promoter $G_{0.82}$ in front of gene VIII and of the strong promoter $G_{0.94}$ which is located within the C-terminal part of gene II (Fig. 1). We now have located yet another weak promoter ($G_{0.01}$) within gene II.

The implications of the presence of several RNA initiation sites within a single gene are by no means clear. It has been reported that the affinity of promoter sites for RNA polymerase can decrease or even be destroyed when the DNA is cleaved in the proximity of these sites (Allett et al., 1974; Maurer, Maniatis and Ptashne, 1974). Furthermore, in the case of T3 DNA, it has been demonstrated that cleavage by restriction enzymes can even result in the artificial creation of RNA initiation sites (Takeya

and Takanami, 1974). Since the RNA species initiated within gene II are made also when intact M13 RF is used as a template, we infer that the promoters ($G_{0.94}$ and $G_{0.01}$) are not *in vitro* artifacts. Promoter $G_{0.94}$ is in fact the strongest RNA polymerase binding site which is located on the M13 genome. Therefore, these promoter sites may have evolved in order to insure that the genes which are located immediately distal to these sites can be expressed frequently. The observation that these promoters are located almost proximal to the 5'-end of the genes (genes V and VIII), the products of which are most abundantly required during the phage infection cycle, is in accord with this assumption. If this is true, then it is still a surprise why these promoters are not located immediately in front of gene V. It is possible that the leading regions, in front of which the promoters $G_{0.01}$ and $G_{0.94}$ are located, code for polypeptides (such as X-protein) which play an essential role in the phage infection cycle. In this connection it is interesting to note that, in case of phage ØX-174, one has also found that, in front of gene B which completely overlaps with the 3'-end of another gene (gene A), a promoter is located (Sanger *et al.*, 1977; Smith *et al.*, 1977). Therefore, the hypothesis that overlapping genes only can be expressed into functional polypeptides when immediately in front of their coding regions a promoter is located deserves further investigation.

Although several restriction fragments used in this study contain the N-terminal part of gene III, our transcriptional results obtained with these fragments have failed to indicate the existence of a promoter in front of this gene. The reason for this failure is unknown, but if such a gene III promoter exists then we have to conclude, under the *in vitro* conditions used, that its affinity for RNA polymerase must be very weak. A low frequency of transcription of gene III is contradictory to the results we obtained from *in vitro* protein synthesis studies. When protein synthesis occurs under the direction of intact M13 RF in a DNA-dependent system, the net synthesis of gene III protein is relatively high (cf. Figs. 3e and 5g; Konings *et al.*, 1975). These observations suggest that additional regulatory mechanisms operate during the expression of gene III. An attractive

hypothesis is that the promoter of gene III is located immediately in front of the central terminator. This, in turn, implies that gene III can be expressed only if termination of transcription is not stringent. This is in fact what we have found. Transcription studies on fragment Hap II-B_2 and Hae II-A have indicated that approximately 10% of the RNA polymerase molecules leak through the termination signal. Also upon transcription of M13 RF several RNA species are found which are greater than one genome in length (Fig. 2e). Since these RNA species are initiated at different promoter sites they are heterogenous in size and, therefore, can be recognized as a diffuse band in the high molecular weight region of the polyacrylamide gel (cf. Fig. 2e). If such a read-through mechanism also operates in the infected cell, then the conclusion that the expression of gene III is regulated at the level of termination of transcription appears to be warranted. If true, the central termination signal has a dual function in the infection cycle: first, as a terminator for the release of most of the initiated RNA chains, and second, as an attenuator which regulates the expression of the genes located immediately distal to this site.

REFERENCES

1. Allet, B., Roberts, R.J., Gesteland, R.F. and Solem, R. (1974). Class of promoter sites for Escherichia coli DNA-dependent RNA polymerase. *Nature*, 249, 217-221.
2. Chan, T., Model, P. and Zinder, N.D. (1975). In vitro protein synthesis directed by separated transcripts of bacteriophage F1 DNA. *J. Mol. Biol.*, 99, 369-382.
3. Denhardt, D. (1975). The single-stranded DNA phages. *CRC Critical Reviews in Microbiology*, pp. 116-123, The Chemical Rubber Company, Cleveland, Ohio.
4. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Physical mapping of the central terminator for transcription on the bacteriophage M13 genome. *Nucleic Acid Res.*, 2, 1811-1820.
5. Edens, L., van Wezenbeek, P., Konings, R.N.H. and Schoenmakers, J.G.G. (1976). Mapping of promoter sites on the genome of bacteriophage M13. *Eur. J. Biochem.*, 70, 577-587.
6. Fellner, P. (1974). Structure of the 16S and 23S ribosomal RNAs. In: *Ribosomes*, (M. Nomura, A. Tissières and P. Lengyel, eds.) pp. 169-191, Cold Spring Harbor Press, New York.
7. Henry, T.J. and Pratt, D. (1969). The proteins of bacteriophage M13. *Proc. Natl. Acad. Sci. U.S.A.*, 62, 800-809.
8. Horiuchi, K., Vovis, G.F., Enea, V. and Zinder, N.D. (1975). Cleavage maps of bacteriophage F1. Location of the Escherichia coli B specific modification sites. *J. Mol. Biol.*, 95, 147-165.
9. Horiuchi, K. and Zinder, N.D. (1976). Origin and direction of synthesis of bacteriophage F1 DNA. *Proc. Natl. Acad. Sci. U.S.A.*, 73, 2341-2345.
10. Jacob, E., Jaenisch, R. and Hofschneider, P.H. (1970). Replication of the small coliphage M13: Evidence for a long-living M13 specific messenger RNA. *Nature*, 277, 59-60.
11. Konings, R.N.H. (1973). Synthesis of phage M13 specific proteins in a DNA-dependent cell-free system. *FEBS Lett.*, 35, 155-160.

12. Konings, R.N.H. and Bloemendal, H. (1969). Synthesis of lens protein in vitro. Ribonucleic acid with template activity isolated from calf lens tissue. *Eur. J. Biochem.*, 7, 165-173.
13. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975). Identification and characterization of the in vitro synthesized products of the bacteriophage M13. *J. Virol.*, 15, 570-584.
14. Konings, R.N.H., Jansen, J., Cuypers, T. and Schoenmakers, J.G.G. (1973). Synthesis of bacteriophage M13-specific proteins in a DNA-dependent cell-free system. II. In vitro synthesis of biologically active gene 5 protein. *J. Virol.*, 12, 1466-1472.
15. Konings, R.N.H., Ward, R., Francke, B. and Hofschneider, P.H. (1970). Gene order of RNA bacteriophage M12. *Nature*, 266, 604-607.
16. Maurer, R., Maniatis, T. and Ptashne, M. (1974). Promoters are in the operators in phage lambda. *Nature*, 249, 221-223.
17. Model, P. and Zinder, N.D. (1974). In vitro synthesis of bacteriophage f1 proteins. *J. Mol. Biol.*, 83, 231-252.
18. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, M. (1975). Studies on bacteriophage fd DNA. II. Localization of RNA initiation sites on the cleavage map of the fd genome. *J.Mol. Biol.*, 95, 33-44.
19. Okamoto, T., Sugiura, M. and Takanami, M. (1969). Length of RNA transcribed on the replicative form DNA of coliphage fd. *J. Mol. Biol.*, 45, 101-111.
20. Ray, D.S. (1977). Replication of filamentous bacteriophages. In: *Comprehensive Virology*. (H. Fraenkel-Conrat and R.R. Wagner, eds.), pp. 105-178. Plenum Publishing Corporation, New York.
21. Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison III, C., Slocombe, P.M. and Smith, M. (1977). Nucleotide sequence of bacteriophage ØX-174 DNA. *Nature*, 265, 687-695.
22. Seeburg, P.H., Nüsslein, L. and Schaller, H. (1977). Interaction of RNA polymerase with promoters from bacterio-

phage fd. *Eur. J. Biochem.*, 74, 107-112.

23. Seeburg, P.H. and Schaller, H. (1975). Mapping and characterization of promoters in phage fd, f1 and M13. *J. Mol. Biol.*, 92, 261-277.
24. Smith, M., Brown, N.L., Air, G.M., Barrell, B.G., Coulson, A.R., Hutchison III, C.A. and Sanger, F. (1977). DNA sequence at the C-termini of the overlapping genes A and B in bacteriophage ØX-174. *Nature*, 265, 702-706.
25. Suggs, S.V. and Ray, D.S. (1977). Replication of bacteriophage M13. XI Localization of the origin of M13 single-strand synthesis. *J. Mol. Biol.*, 110, 147-163.
26. Sugimoto, K., Okamoto, T., Sugisaki, H. and Takanami, M. (1975). The nucleotide sequences of an RNA polymerase binding site on bacteriophage fd DNA. *Nature*, 253, 410-414.
27. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M.(1977). Studies on bacteriophage fd DNA. IV. The sequence of messenger RNA for the Major Coat Protein Gene. *J. Mol. Biol.*, 110, 487-507.
28. Takeya, T. and Takanami, M. (1974). Isolation of DNA segments containing promoters form bacteriophage T3 DNA. *Biochemistry*, 13, 5388-5394.
29. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Regulation of gene activity in bacteriophage M13 DNA: Coupled transcription and translation of purified genes and gene-fragments. *Virology*, 67, 487-497.
30. Van den Hondel, C.A., Pennings, L. and Schoenmakers, J.G.G. (1976). Restriction enzyme cleavage maps of bacteriophage M13: Existence of an intergenic region. *Eur. J. Biochem.*, 68, 55-70.
31. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1973). Cleavage of bacteriophage M13 DNA by Haemophilus influenzae Endonuclease. *Mol. Biol. Rep.*, 1, 41-45.
32. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1975). Studies on bacteriophage M13 DNA. I. A cleavage map of the M13 genome. *Eur. J. Biochem.*, 53, 547-558.
33. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1976). Cleavage maps of the filamentous bacteriophages M13, fd, f1 and ZJ_2. *J. Virol.*, 18, 1024-1039.

34. Van den Hondel, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Studies on bacteriophage M13 DNA. II. The gene order of the M13 genome. *Eur. J. Biochem.*, 53, 559-567.

35. Vovis, G.F., Horiuchi, K. and Zinder, N.D. (1975). Endonuclease R. Eco R II restriction of bacteriophage f1 DNA in vitro: ordering of gene V and VII, location of an RNA promoter for gene VIII. *J. Virol.*, 16, 674-684.

CHAPTER VI

THE NUCLEOTIDE SEQUENCE OF THE CENTRAL TERMINATION SIGNAL FOR TRANSCRIPTION ON THE BACTERIOPHAGE M13 GENOME

THE NUCLEOTIDE SEQUENCE OF THE CENTRAL TERMINATION SIGNAL FOR TRANSCRIPTION ON THE BACTERIOPHAGE M13 GENOME

ABSTRACT

The 8S RNA species is one of the smallest RNA molecules formed upon the *in vitro* transcription of the phage M13 replicative form DNA. This transcript is initiated at the promoter $G_{0.82}$ and is terminated at the central termination signal distal to gene VIII. By establishing the 3'-terminal nucleotide sequence of the 8S RNA species, the structure of the central termination signal for transcription was elucidated. The 3'-OH terminal region of the 8S RNA molecule was isolated as a partially RNase T_1-resistant 'core' fragment of the sequence - AUACAAUUAAAGGCUCC-UUUUGGAGCCUUUUUUUU$_{OH}$. The features of this sequence and their putative roles in the transcription termination event are discussed.

INTRODUCTION

Since initiation and termination of transcription occurs at well-defined regions on DNA genomes, both promoter and termination signals must exhibit specific features which influence the transcription process. In the past few years essential information about the nature of these signals has been provided by studies in which their nucleotide sequences were determined. To elucidate the structural features of a termination signal, usually the 3'-OH terminal ends of RNA molecules have been sequenced (1, 2). Since it is generally assumed that these specific regions of RNA molecules are transcribed from at least part of the DNA regions encompassing termination signals, the sequence information obtained can be used as a clue to the structure of these signals. From nucleotide sequence studies on various RNA transcripts there is suggestive evidence that both G-C and A-T rich blocks are involved in the termination process. Furthermore, RNA sequences transcribed from termination signals often have the intrinsic capacity to form a tight, hairpin-like structure.

In vitro transcription studies performed on the genome of the filamentous phage M13 have indicated that there are several promoter sites but only a single, rho-independent, termination site operative on this template molecule (for a recent review

see 3). The latter, so called 'central' termination site, is located immediately distal to gene VIII and has an important function in establishing the gene expression of this phage genome (4, 5, 6). As a result of the transcription mechanism, the genes which are located immediately upstream of the central termination signal are transcribed very frequently whereas the transcription frequencies of the genes downstream of this signal are low (7, 4, 5). Another consequence of this transcription mechanism is that all *in vitro* synthesized RNA species have identical 3'-OH terminal nucleotide sequences and hence, any of the *in vitro* generated RNA species can be used for structure analysis of the central termination signal.

In this chapter we present the nucleotide sequence of the 3'-OH terminal end of the smallest RNA species formed upon transcription of M13 replicative form DNA (RF) with *E.coli* RNA polymerase holoenzyme. This particular RNA species, which has a length of about 370 nucleotides (8S), is initiated at the strong promoter ($G_{0.82}$) in front of gene VIII and is terminated within a distance of 60 base pairs from the last triplet of this gene. The fragment containing the 3'-OH terminal end of the 8S RNA molecule was isolated as described by Sugimoto *et al.* (8) and its nucleotide sequence was determined by using conventional RNA sequencing techniques.

MATERIALS AND METHODS

Materials

Escherichia coli RNA polymerase holoenzyme was donated by Dr. R. van Meteren. DBAE cellulose was kindly made available by Dr. M. Rosenberg. The restriction endonuclease from *Haemophilus aphirophilus* (endo R. *Hap* II) was isolated as described previously (9). RNase T1 was obtained from Calbiochem, La Yolla, California and bacterial alkaline phosphatase from Boehringer, Mannheim. Pancreatic RNase A, spleen phosphodiesterase and snake venom phosphodiesterase were purchased from Worthington Biochemical Corp., Freehold, New Jersey.

Replicative Form I DNA and Restriction Fragments

The method for the preparative isolation and purification of M13 replicative form I DNA has been described elsewhere (10). The digestion of M13 RF with the restriction enzyme endo R. Hap II, the separation and the subsequent purification of the fragments generated were carried out as described by Van den Hondel et al. (9, 11).

RNA Synthesis in vitro and Isolation of 8S RNA Transcripts

RNA synthesis in vitro was performed in a standard reaction mixture (0.1 ml) which contained: 4 µmol of Tris-HCl (pH 7.9), 15 µmol of KCl, 0.8 µmol of $MgCl_2$, 0.1 µmol of dithiotreitol, 0.01 µmol of EDTA, 0.1% Tween-80, 25 µg of gelatin, 0.2 pmol of M13 DNA restriction fragment Hap II-B_2 and 10.0 pmol of E.coli RNA polymerase holoenzyme. After a pre-incubation period of 3 min at 37°C, ribonucleoside triphosphates were added to a final concentration of 80 µM, except for the (α^{32}P)-labeled triphosphate whose concentration was 8 µM. Per incubation of 0.1 ml, about 10 µCi of the (α^{32}P)-labeled triphosphate (Amersham, spec. act > 100 Ci/mmole) was added. After a further incubation for 20 min, the reaction was terminated by the addition of 10 µl of 10% SDS and 0.1 ml of phenol saturated with water. Prior to alcohol precipitation 10 µg of carrier tRNA was added. After two precipitations the RNA was dissolved in 20 µl of deionized formamide.

The (α^{32}P)-labeled 8S RNA product was purified by electrophoresis through 3.5% polyacrylamide gels under denaturing conditions (5). The position of the 8S RNA species was detected by autoradiography after which the band containing the radioactively-labeled RNA product was sliced from the gel. The RNA was eluted by repeated extraction at room temperature for 2 hrs in a buffer containing 0.1 M NaCl, 0.01 MTris -HCl (pH 7.6), 0.001 M EDTA and 0.1% SDS. After low speed centrifugation, the supernatant was collected and the RNA was precipitated by the addition of 20 µg of carrier RNA, 0.1 vol. of 3M sodium acetate (pH 5.3) and 2.5 volume of cold ethanol.

Isolation of 3'-OH terminal ends of 8S RNA

i. Isolation by polyacrylamide gelelectrophoresis

The 3'-OH terminal end of the 8S RNA species was isolated by cleavage of the purified transcript with RNase T1 at an enzyme to substrate ratio of 1 : 10 (w/w) and subsequent fractionation of the RNA fragments by electrophoresis through 12.5% acrylamide gels in the presence of 7M urea. Recovery of the partially RNase T1-resistant 'core' fragment from the acrylamide gel was performed essentially as described for the 8S RNA species.

ii. Isolation by DBAE-column chromatography

The isolation of 3'-OH terminal oligonucleotides by DBAE column chromatography was carried out according to the procedure described by Rosenberg et al. (12). The RNase T1-digestion products of 8S mRNA were chromatographed on a column (30 x 0.5 cm) of DBAE cellulose and oligonucleotides with 3'-terminal phosphate moieties were eluted from the column with 0.05 M morpholinum chloride, 0.1 M $MgCl_2$, 1.0 M NaCl, 20% dimethylsulfoxide, pH 8.7 (solvent A). Under these conditions oligonucleotides containing 3'-OH groups are selectively retained and can be subsequently eluted by solvent A containing 0.1 M sorbitol. The 3'-OH terminal oligonucleotides are then desalted by dialysis for 15 hours against three changes of 6 liters each of distilled water and finally concentrated by freeze-drying.

Nucleotide sequence analysis of the 3'-OH terminal RNA fragment

The nucleotide sequence of the purified RNase T1-resistant 'core' fragment was determined according to the methods described by Sanger and co-workers (13,14). The 'core' fragment was hydrolysed to completion either by treatment with RNase T1 or by treatment with pancreatic RNase. RNase T1 digestion was carried out for 30 min at 37°C with an enzyme to substrate ratio of 1 : 1 (w/w). Pancreatic RNase digestion was carried out for 60 min at 37°C with an enzyme to substrate ratio of 1 : 10 (w/w). Oligonucleotides resulting from either RNase T1 digestions or pancreatic RNase digestions were separated by electrophoresis on cellulose acetate strips in 5% (v/v) acetic acid - 7M urea,

adjusted to pH 3.5 with pyridine (14). After electrophoresis the products were transferred to DEAE thin-layer sheets by blotting and further separated by homochromatography at 65°C with homomixture c (13). Oligonucleotides from the fingerprints were further characterised by digestion with other nucleases (e.g. RNase T1, pancreatic RNase A, snake venom phosphodiesterase and spleen phosphodiesterase) and by alkaline hydrolysis. Partial digestions with snake venom phosphodiesterase and spleen phosphodiesterase were carried out as described by Brownlee (15) after which the resulting oligonucleotides were fractionated by two-dimensional homochromatography. The nucleotide sequences were deduced from the characteristic mobility shifts of the partially hydrolysed products.

RESULTS

Digestion of phage M13 RF with endonuclease Hap II results in the formation of several DNA fragments which have been ordered in a circular restriction enzyme cleavage map (9). One of the fragments generated, designated Hap II-B_2, contains the central termination signal and a strong promoter $G_{0.82}$ at about 370 nucleotides upstream of this signal (4). Thus, for structure analysis of the central termination signal highly labeled 8S RNA transcripts can easily be obtained by using fragment Hap II-B_2 as a template for transcription.

In an initial attempt to isolate 3'-OH terminal oligonucleotides of the 8S RNA molecule, the purified transcripts were digested with RNase T1 and the digestion products were chromatographed on a column of dihydroxyboryl substituted cellulose (DBAE; 16). This material, which selectively binds oligonucleotides with vicinal hydroxyl groups, has proven to be very useful for the isolation of 3'-OH terminal oligonucleotides from a variety of RNA species (16, 12, 17). In control experiments with uniformly (α^{32}P)-labeled 16S ribosomal RNA as a probe, this technique enabled us to isolate selectively the 3'-OH containing RNA fragment. However, for reasons still unknown, we failed to isolate the 3'-OH terminal oligonucleotide of the digested 8S RNA species with this procedure. To circumvent this unexpected

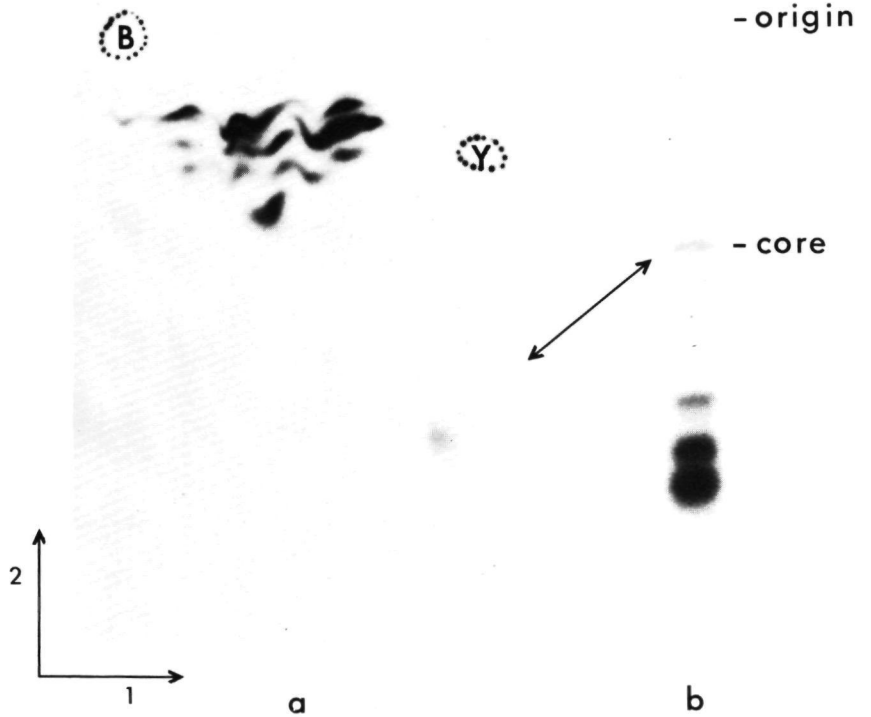


Fig. 1: Fractionation of products resulting from limited RNase T1 digestion of the purified 8S RNA species. Following digestion of the 8S RNA with RNase T1 (1 : 10, w/w, enzyme to RNA) the RNase resistant 'core' was separated from the other digestion products by means of (a) electrophoresis on cellulose-acetate (pH 3.5) in the first dimension (arrow 1) followed by homochromatography on DEAE thin-layer sheets using homomixture b in the second dimension (arrow 2). (b) Electrophoresis through 12.5% polyacrylamide gels containing 7M urea.

difficulty, we decided to isolate the 3'-OH terminal end of the 8S RNA species by an alternative procedure described by Sugimoto et al. (8). In this procedure use was made of the ability of several 3'-terminal RNA regions to form stable, hairpin-like, secondary structures. Since RNase T1 exhibits a preference for single-stranded RNA molecules, these structures are less susceptible to digestion with this enzyme and can, therefore, be isolated as RNase T1-resistant 'core' fragments (18, 14, 23). Digestion of the phage M13 specific 8S RNA species with RNase T1

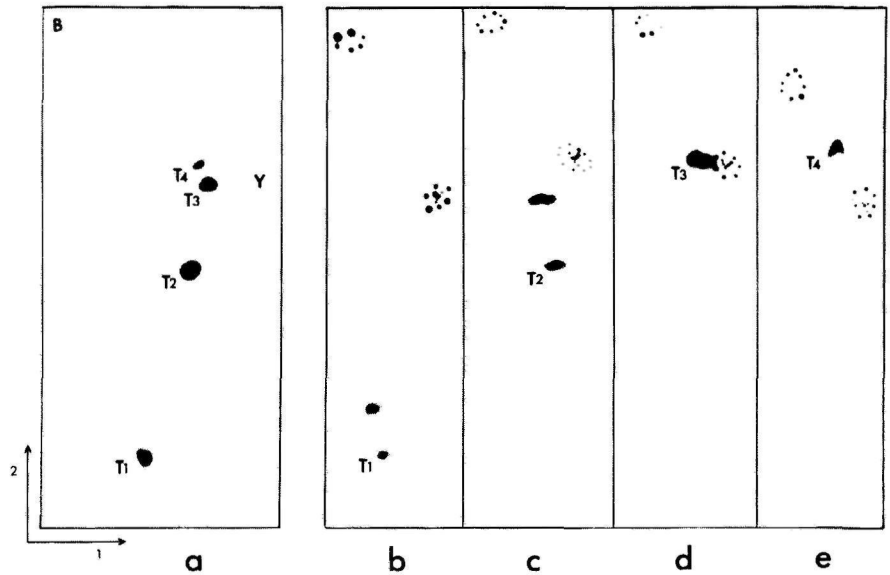


Fig. 2: Effect of alkaline phosphatase treatment on the mobility of isolated T1-oligonucleotides obtained after exhaustive digestion of the (α^{32}P) UTP-labeled 'core' fragment. The 'core' fragment was subjected to exhaustive digestion with RNase T1 (1 : 1, w/w, enzyme to RNA) and the resulting oligonucleotides were fractionated by means of two-dimensional homochromatography using homomixture c (plate a). Each T1-oligonucleotide was isolated and divided into two equal parts. One part was treated with bacterial alkaline phosphatase and, after termination of the reaction, mixed with the untreated sample. This mixture was again subjected to the fingerprint procedure (plates b-e).

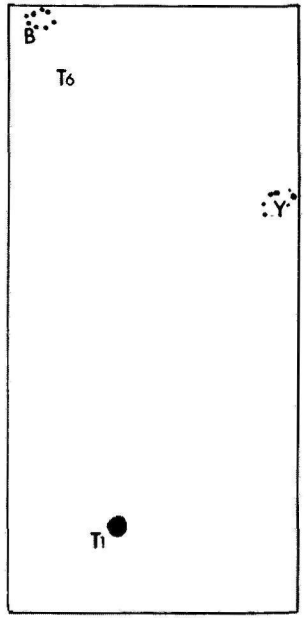
under standard conditions (1 : 10, w/w, enzyme to RNA) also results in the formation of a relatively large RNase T1-resistant fragment. Due to its size, this fragment can readily be separated from the other oligonucleotides either by homochromatography (Fig. 1a) or by electrophoresis through 12.5% acrylamide gels (Fig. 1b). In view of its mobility on acrylamide gels the size of this 'core' fragment is about 40 nucleotides.

If this large RNA fragment represents the 3'-terminal end of the 8S RNA species, then one expects that upon incubation of this fragment with relatively high amounts of RNase T1 it will be cleaved into a number of smaller T1-oligonucleotides (8). Additionally, it may be expected that one of the newly generated oligonucleotides represents the ultimate 3'-OH terminal end of

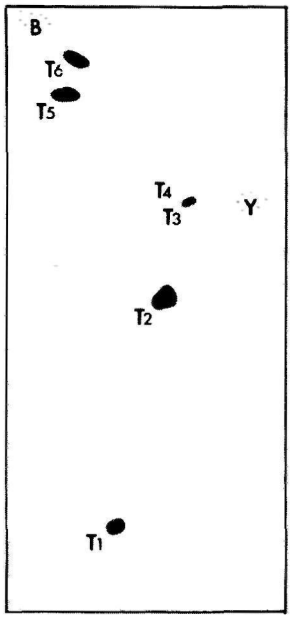
the 8S RNA species. To test this hypothesis, the isolated 'core' fragment was incubated with RNase T1 under vigorous digestion conditions (1 : 1, w/w, enzyme to RNA) and the digest was analysed by two-dimensional homochromatography. As shown in Fig. 3, six T1-oligonucleotides are now apparent. Each T1-oligonucleotide obtained was isolated from the thin-layer sheet and subsequently treated with bacterial alkaline phosphatase. Since a loss of phosphate group can be traced due to the increased mobility of the T1-oligonucleotide upon homochromatography, the presence of oligonucleotides containing free 3'-OH groups can be readily detected. With the aid of this technique it was demonstrated unambigiously that the oligonucleotides T3 and T4 represent the ultimate 3'-OH terminal end of the 8S RNA species. (Fig. 2).

To establish the nucleotide sequence of the PNaseT1-resistant 'core' fragment by nearest neighbour analysis, the 8S RNA was synthesized in the presence of each one of the four (α^{32}P)-labeled ribonucleoside triphosphates and purified by electrophoresis through a 3.5% acrylamide gel. The isolated 8S RNA was subjected to limited digestion with RNase T1 (1 : 10, w/w, enzyme to RNA) and the resulting digestion products were fractionated by electrophoresis through 12.5% acrylamide gels. The largest product (i.e. the 'core' fragment) was isolated from the acrylamide gel and subsequently digested to completion by incubation with RNase T1 (1 : 1, w/w, enzyme to RNA). The products of this reaction were analysed by two-dimensional homochromatography (13). The separated T1-oligonucleotides (Fig. 3) were isolated and characterised by analysis of the products obtained after pancreatic RNase digestion and alkaline hydrolysis. The data obtained by nearest neighbour analysis were sufficient to deduce unambiguously the nucleotide sequences for almost all T1-oligonucleotides. Two exceptions were the oligonucleotides T1 and T2 (Fig. 3). Their nucleotide sequences were confirmed by partial digestions with snake venom phosphodiesterase and spleen phosphodiesterase.

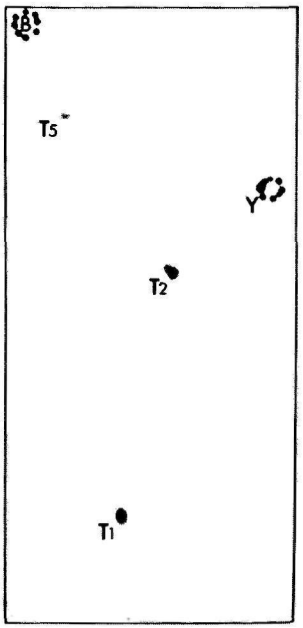
Digestion of the 'core' fragment with pancreatic RNase (1 : 10, w/w, enzyme to RNA) yielded several oligonucleotides of which only the two largest (P1 and P2; Fig. 4) were isolated. Their respective nucleotide sequences were deduced by RNase T1 digestion and alkaline hydrolysis (Table 1). Since these two pancreatic products contain all guadenylate residues of the 'core'



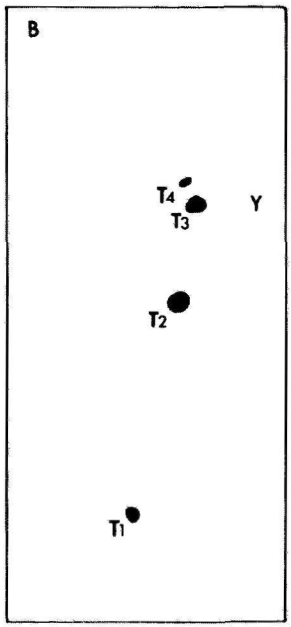
r-ATP



r-CTP



r-GTP



r-UTP

T_1-Products					
	GTP	ATP	UTP	CTP	deduced sequence
1	AAAG	AAAG, AAU AU; AC, U	AAU; AU	AC	AUACAAUUAAAG
2	G, U	–	3U, 2C	U; C	CUCCUUUUG
3	–	–	7U, C	C	$CCUUUUUUUU_{OH}$
4	–	–	6U, C	C	$CCUUUUUUU_{OH}$
5	AG	–	–	AG	AG
6	–	G	–	G	G
Large Panc. Products					
1	AAAG	AAAG	C	G	AAAGGC
2	AG, G	G	–	AG, C	GGAGC
Complete sequence of the core-fragment: $-AUACAAUUAAAG.G\ CUCCUUUUG\ G.AG.CCUUUUUUUU_{OH}$					

Table 1: Composition of T1-oligonucleotides and large pancreatic products as obtained from the 'core' fragment. Oligonucleotide fragments (Fig. 3 and Fig. 4) were eluted from the DEAE cellulose thin-layer sheet and characterised by digestion with either pancreatic RNase or RNase T1 and subsequently analysed by electrophoresis on DEAE paper at pH 3.5. The component products were further analysed by alkaline hydrolysis. The radioactively labeled cleavage products of the various T1-oligonucleotides and of the two large pancreatic oligonucleotides are shown as a function of the (α^{32}P)-labeled precursor which has been incorporated into the RNA. The distribution of radioactive phosphate groups in the labeled products is indicated by underlining. The sequences of the oligonucleotides T1 and T2 were obtained by partial digestions with snake venom phosphodiesterase. The 5'-terminal sequence of oligonucleotide T1 was established by partial digestion with spleen phosphodiesterase.

Fig. 3: Two dimensional fractionation of oligonucleotides obtained upon exhaustive digestion of the 'core' fragment with RNase T1. Following digestion of the 'core' fragment with RNase T1 (1 : 1, w/w, enzyme to RNA), the oligonucleotides were separated by electrophoresis on cellulose-acetate (pH 3.5) in the first dimension (arrow 1) and by homochromatography on DEAE thin-layer sheets using homomixture c in the second dimension (arrow 2). The suffixes ATP, CTP, GTP and UTP refer to the respective (α^{32}P)-labeled precursors which have been used for radioactive labeling of the 8S RNA species. Each of the generated T1-oligonucleotide products has been assigned a number (T1-T6) which corresponds to the number used in Table 1.

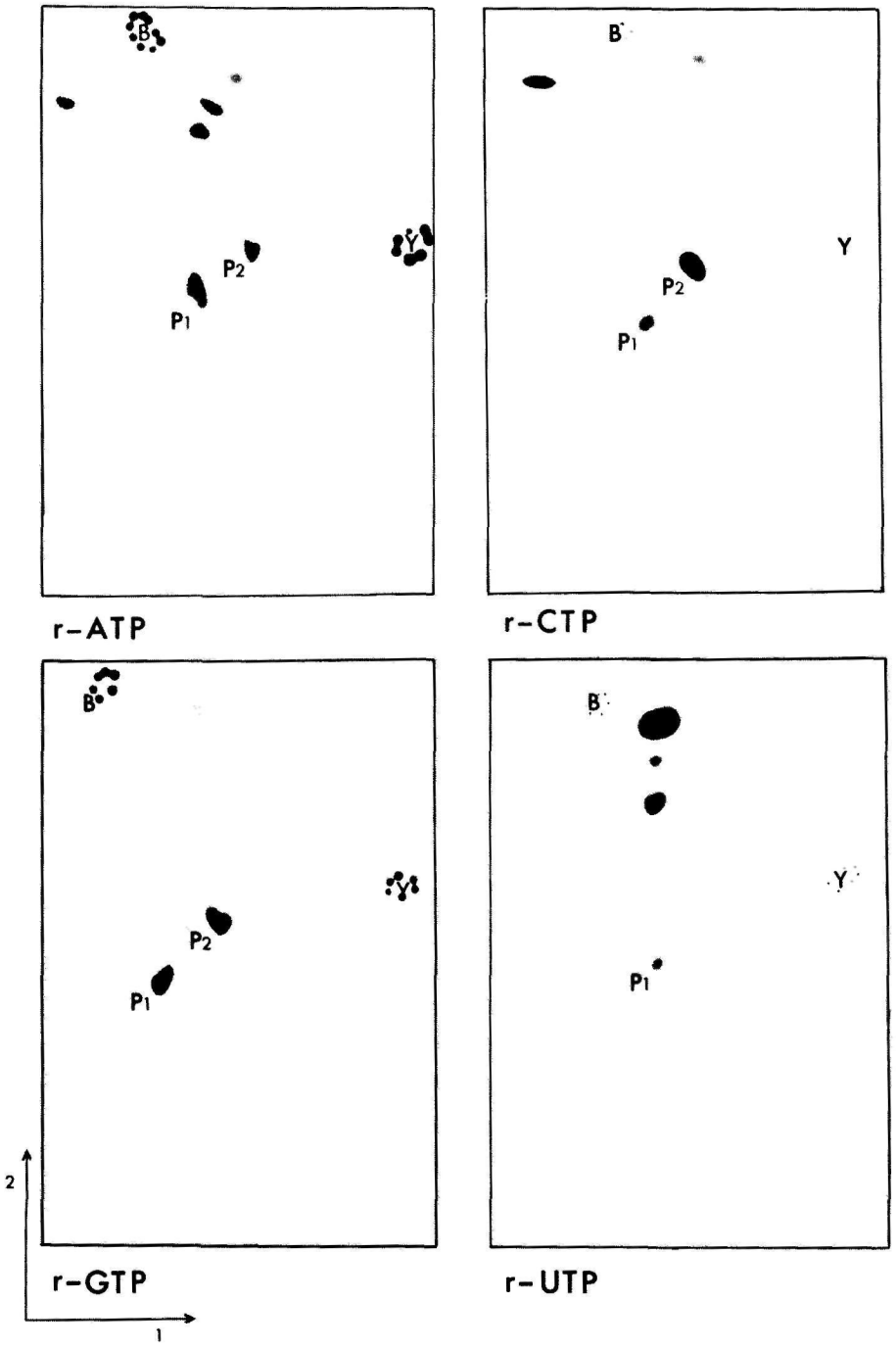

B
P2
P1
Y
r-ATP
B
Y
P2
P1
r-CTP
B
P2
P1
r-GTP
2
1
B
Y
P1
r-UTP

Fig. 4: Two-dimensional fractionation of oligonucleotides obtained upon digestion of the 'core' fragment with pancreatic RNase A. Following digestion of the 'core' fragment with RNase A (1 : 10, w/w, enzyme to RNA), the oligonucleotides were separated by electrophoresis on cellulose-acetate (pH 3.5) in the first dimension (arrow 1) and by homochromatrography on DEAE thin-layer sheets using homomixture c in the second dimension (arrow 2). The suffixes ATP, CTP, GTP and UTP refer to the respective (α^{32}P)-labeled precursors which have been used for radioactive labeling of the 8S RNA species. The two largest pancreatic products have been assigned a number (P1 and P2) which corresponds to the number used in Table 1.

fragment, their nucleotide sequences greatly facilitated the ordering of the six RNase T1-oligonucleotides. For example, the sequence GGAGC permits an unambiguous ordering of the fragments T2, T6 and T3 (T4). In the same way the sequence AAGGC provided the overlap for the fragments T1, T6 and T2 (Table 1). From these arrangements it was deduced that the oligonucleotide T3 (T4) represents the 3'-OH terminal end of the 'core' fragment. An identical conclusion was reached in the previous paragraph based upon the results obtained with bacterial alkaline phosphatase (Fig. 2). Both results suggest, therefore, that the *in vitro* synthesized 8S mRNA species is terminated in an A-T rich region of the phage genome after either the seventh or eighth adenylate residue. Based upon the molar relative yields of the oligonucleotides T3 and T4 (Fig. 3), it may be concluded that, under the *in vitro* conditions used, RNA synthesis is terminated preferentially (more than 90%) after the eighth adenylate residue.

The nucleotide sequence data obtained show that none of the five guadenylate residues present within the 'core' fragment is susceptible to cleavage with RNase T1 under standard digestion conditions. An explanation for this observation is that the five guadenylate residues form part of a region within the 8S RNA species with a tight secondary structure. A similar explanation is strongly supported by the nucleotide sequence derived which indicates that part of the 'core' fragment has the intrinsic capacity to form a hairpin-like structure (Fig. 5). Since a completely identical loop structure has recently been proposed for the 3'-OH terminal region of the 8S RNA species transcribed from phage fd DNA (8) , it is clear that for both phages these termination signals are identical.

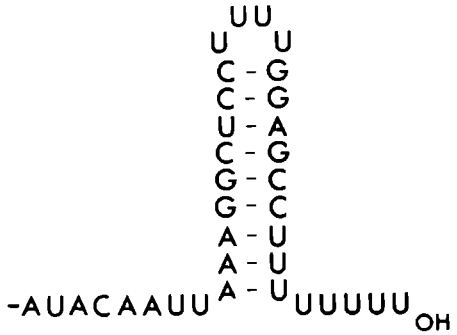


Fig. 5: Possible secondary structure for the 3'-terminal region of the 8S RNA species as deduced from the nucleotide sequence of the 'core' fragment.

DISCUSSION

Several of the RNA products which are synthesized in vitro by E.coli RNA polymerase share a common arrangement of nucleotides at their 3'-ends. The similarities include the presence of G-C rich blocks (four to ten nucleotides) at approximately eight to ten nucleotides upstream of a string to six to eight uridenylate residues. In addition, the 3'-OH terminal regions of RNA molecules often have the intrinsic property of forming hairpin-like structures (2) . Although the function of these features in the termination proces is not yet clear, it has been noted by Maizels (20) that G-C rich regions in the *lac* operon are able to cause a 'pausing' in the transcription process. Furthermore, it was observed that base-pair changes altering these G-C rich regions influence the 'pausing' event(21). Based on these data, Gilbert (1) has suggested that the function of the G-C rich region on the template DNA is to block the motion of the RNA polymerase and that the string of adenylate residues represents a weak hydrogen-bonding region that enhances the ability of the RNA molecule to dissociate from the DNA. The relevance of the length of the adenylate-rich region on the template molecule in the termination event is indicated by the observation that a deletion removing the distal four of the series of eigth A-T base-pairs in the E.coli *trp* attenuator abolishes (at least

in vitro) the termination of transcription at this site (19). On the other hand, the stability of the base-paired loop structure in the 3'-OH terminal region of RNA molecules is possibly an additional factor which influences the termination phenomenon. Substituting the analog ITP (inosine triphosphate) for GTP in the in vitro transcription system resulted in the synthesis of RNA molecules in which the G-C base pairs formed in the hairpin structure are replaced by the weaker interacting I-C base pairs. As a result of this substitution, a reduced termination efficiency was observed (19). Similar observations were made for the efficiency of a transcription termination signal on the genome of bacteriophage lambda (22). In the latter study it was observed that single-base changes increasing the stability of the loop structure in the RNA are accompanied by an enhanced termination efficiency. Single-base changes destabilising this loop structure were found to reduce the efficiency of the termination event. These data strongly suggest that weaker binding forces in the hairpin-like structure lower the efficiency with which RNA synthesis is terminated. A possible explanation for this phenomenon is, that due to the base-pairing capacity of the hairpin-like structure, the growing RNA chain is suddenly withdrawn from the active site of the RNA polymerase molecule causing termination of transcription. An alternative possibility is that the secondary structure formed in the RNA molecule intervenes with the RNA polymerase molecule in such a way that the chain elongation centre is inactivated. To what extent RNA polymerase is involved in the subsequent process of release of the RNA molecule from the tertiary complex is unknown.

According to the data presented in this paper, the central termination signal on the phage M13 genome shares several features with transcription termination signals described for other organisms. That is, the central termination signal contains blocks rich in G-C and A-T base-pairs and the RNA transcribed from this signal has the intrinsic property to form a tight, hairpin-like, structure. The observation that the termination events do not occur at a single, well-defined site on the DNA molecule but at several consecutive nucleotides is also in agreement with data obtained in other in vitro experiments (23, 22, 8). On the

M13 genome these termination events occur after transcription of the seventh or eighth adenylate residue of the termination signal. The yields of the 3'-OH containing T1-oligonucleotides of the 'core' fragment indicate that termination events at the seventh adenylate residue are rare (less than 10%). A similar observation has been made for the *in vivo* termination of 8S RNA. Recent data obtained by Rivera *et al.*(24), show that the 'core' fragment of the 8S RNA isolated from phage M13 infected *E.coli* cells is also predominantly terminated at the eighth adenylate residue. It must therefore be concluded that the 3'-OH terminal heterogeneities are inherent in the transcription termination mechanism and are not the result of the *in vitro* transcription conditions. Based upon this observation and the observation that the 3'-OH terminal nucleotide sequences of the *in vitro* and *in vivo* generated 8S RNA species are completely identical, it can further be concluded that termination of transcription occurs with high specificity under *in vitro* conditions.

REFERENCES

1. Gilbert, W. (1976). In: *RNA Polymerase*. pp. 193-205 (R. Losick and M. Chamberlain, eds.), Cold Spring Harbor Laboratory.
2. Roberts, J.W. (1976). In: *RNA Polymerase*. pp. 247-271 (R. Losick and M. Chamberlain, eds.), Cold Spring Harbor Laboratory.
3. Konings, R.N.H. and Schoenmakers, J.G.G. (1978). In: *Single-Stranded DNA Phages*. (D.T. Denhardt, D.T. Ray and D. Dressler, eds.), Cold Spring Harbor Laboratory, in press.
4. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Nucleic Acid Res.*, 2, 1811-1820.
5. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). *Virology*, 86, 354-367.
6. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). *J. Virol.*, in press.
7. Chan, T., Model, P. and Zinder, N.D. (1975). *J. Mol. Biol.*, 99, 369-382.
8. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1977). *J. Mol. Biol.*, 110, 487-507.
9. Van den Hondel, C.A.M. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 547-558.
10. Van den Hondel, C.A.M. and Schoenmakers, J.G.G. (1976). *J. Virol.*, 18, 1024-1039.
11. Van den Hondel, C.A.M., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 559-567.
12. Rosenberg, M., Weissman, S. and de Crombrugghe, B. (1975). *J. Biol. Chem.*, 250, 4755-4764.
13. Brownlee, G.G. and Sanger, F. (1969). *Eur. J. Biochem.*, 11, 395-399.
14. Sanger, F., Brownlee, G.G. and Barrel, B.G. (1965). *J. Mol. Biol.*, 13, 373-379.
15. Brownlee, G.G. (1972). *Laboratory Techniques in Biochemistry and Molecular Biology: Determination of Sequences in RNA*.(T.S. Work and E. Work, eds.) North-Holland Publishing Company.

16. Rosenberg, M. (1974). *Nucleic Acid Res.*, 1, 653-671.
17. Sprague, K.U. and Steitz, J.A. (1975). *Nucleic Acid Res.*, 2, 787-798.
18. Guilley, H., Jonard, G. and Hirth, L. (1975). *Proc. Natl. Acad. Sci. U.S.A.*, 72, 864-868.
19. Lee, F. and Yanofsky, C. (1977). *Proc. Natl. Acad. Sci. U.S.A.*, 74, 4365-4369.
20. Maizels, N.M. (1973). *Proc. Natl. Acad. Sci. U.S.A.*, 70, 3585-3589.
21. Gilbert, W., Maizels, N.M. and Maxam, A. (1974). *Cold Spring Harbor Symp. Quant. Biol.*, 38, 845-856.
22. Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978). *Nature*, 272, 414-423.
23. Lee, F., Squires, C.L., Squires, C. and Yanofsky, C. (1976). *J. Mol. Biol.*, 103, 383-393.
24. Rivera, M.J., Smits, M.A., Quint, W., Schoenmakers, J.G.G. and Konings, R.N.H. (1978). *Nucleic Acid Res.*, 3, 2895-2912.

CHAPTER VII

TRANSCRIPTION OF BACTERIOPHAGE M13 DNA
EXISTENCE OF PROMOTERS DIRECTLY
PRECEDING GENES III, VI AND I

TRANSCRIPTION OF BACTERIOPHAGE M13 DNA
EXISTENCE OF PROMOTERS DIRECTLY PRECEDING GENES III, VI AND I

ABSTRACT

In vitro transcription and coupled transcription-translation studies have been performed with restriction fragments of M13 replicative form DNA which contain either gene III, gene VI or gene I. It could be demonstrated that DNA fragments which contain only gene III were able to direct the synthesis of gene III protein. Fragments which encompassed genes VI and I gave rise to the synthesis of gene I protein only whereas gene I-containing fragments were able to direct the synthesis of gene I protein. None of the fragments studied gave rise to a detectable level of gene VI protein, although an RNA transcript of gene VI could readily be obtained during *in vitro* transcription of the relevant gene VI-containing DNA fragments. From these results we have concluded that the promoters $A_{0.56}$ and $A_{0.51}$ are located in front of gene VI and I, respectively, and that gene III is also equipped with a promoter ($X_{0.75}$).

Introduction of a single cleavage within the gene III region does not abolish the expression of genes VI and I *in vitro*. Hence, the expression of these genes is not solely dependent on the initiation of RNA synthesis at the gene III promoter or on leakage of transcription through the central termination site ($T_{0.75}$) but is also determined by the initiation frequency of RNA synthesis at their individual promoters.

INTRODUCTION

Control of transcription requires that initiation as well as termination occurs at appropriate sites on the DNA template. For a better understanding of this process, a detailed knowledge of the positions and respective efficiencies of these sites is essential.

During the last few years the *in vitro* transcription process of the filamentous phages has been studied extensively (for reviews see 2, 17). From the results of these studies we know now that transcription of the M13 genome starts at eight promoter sites and that termination of transcription occurs at only a single site. Three of these promoters initiate the synthesis of RNA chains which start with pppA (A-promoters) whereas the other five promoters initiate RNA synthesis which starts with pppG (G-promoters) (4, 13, 18). It has also been demonstrated that

the G-promoters are located immediately in front of the genes VIII, V, X and II (1, 4, 5, 25). Of the A-promoters only one has been positioned so far on the genetic map of phage M13. By transcription and translation experiments it has been shown unambiguously that this A-promoter is located immediately in front of gene IV (10, 25). The genetic position of the two remaining A-promoters is not known, although from transcription data their positions on the physical map have been calculated (4). Their calculated positions suggest a location immediately in front of the genes VI and I. Direct proof for this assumption, however, is still lacking.

In the absence of rho-factor, only one strong termination signal operates on the M13 genome (3, 5, 13). This central termination site has been localized in the intercistronic region between the genes VIII and III (1, 3, 5, 21). Under in vitro conditions the termination of transcription at this site is not stringent. Approximately 10% of the RNA polymerase molecules leak through the central termination site and enhance the transcription frequency of the genes which are located distal to this site (4-6). Since up to now transcription studies have not given evidence for the existence of a strong promoter in front of gene III (4, 5, 13, 18), it has been postulated that the expression of gene III is accomplished by this read-through effect. However, the existence of a promoter in front of gene III with a very weak initiation capacity could not be excluded (1, 7, 13).

In this study we have focussed our attention to the question whether or not the genes III, VI and I are preceded by promoter sites. Therefore, we have isolated several restriction fragments which contained either the complete gene III, gene VI or gene I and have analysed the template function of each fragment in both transcription and translation experiments. Upon translation of the DNA fragments in a DNA-dependent protein synthesizing system, the synthesis of the proteins encoded by both gene III and gene I could be demonstrated. From these results we have concluded that these genes are preceded by promoter sites. A similar conclusion was reached from the results of transcription studies of the relevant DNA fragments.

MATERIALS AND METHODS

Enzymes

Escherichia coli RNA polymerase holoenzyme was a generous gift from Dr. R. van Meteren (University of Leiden). The restriction endonuclease from Arthrobacter luteus (endoR.Alu I) was a kind gift from Dr. P. Baas (University of Utrecht). The restriction endonucleases from Haemophilus aegyptius (endoR. Hae II and endoR. Hae III) have been isolated as described previously (23, 26). The restriction endonuclease endoR. BamH I was obtained from Boehringer, Germany.

Replicative form I DNA and restriction fragments

The method for preparative isolation and purification of M13 replicative form I DNA has been described (23). Digestion of M13 RF with the restriction endonucleases has been performed as described by Van den Hondel and Schoenmakers (22, 23). After digestion, the fragments were separated by electrophoresis on 2.5% (for endoR. Hae II and endoR. Alu I fragments) or on 3.0% discontinuous polyacrylamide slab gels and further purified by sucrose density gradient centrifugation as described by Van den Hondel et al., (24).

RNA synthesis in vitro

RNA synthesis in vitro was performed in a standard reaction mixture (0.1 ml) which contained: 4 µmol of Tris-HCl (pH 7.9), 15 µmol of KCl, 0.8 µmol of $MgCl_2$, 0.1 µmol of dithiothreitol, 0.01 µmol of EDTA, 0.1 µmol of UMP, 0.1% Tween 80, 25 µg of bovine serum albumin, 0.2 pmol of M13 DNA restriction fragment and 8.0 pmol of E.coli RNA polymerase holoenzyme. After a preincubation period of 5 min at 37°C, ribonucleoside triphosphates were added to a final concentration of 80 µM, except for the (α-^{32}P) CTP the concentration of which was 8 µM. After further incubation for 10 min the reactions were terminated by the addition of 10 µl of 10% SDS and 0.1 ml of phenol saturated with a buffer containing 0.01 M Tris-HCl (pH 7.6) and 0.001 M EDTA.

After RNA synthesis, 10 µg of carrier tRNA was added and the reaction mixtures were extracted with phenol. Subsequently, the RNA was precipitated twice with 2.5 volumes of cold ethanol and finally dissolved in 15 µl of deionized formamide. The RNA products synthesized *in vitro* were then analysed on 1.9% polyacrylamide slab gels (50 cm x 20 cm x 0.15 cm) in the presence of 7M urea (5).

Protein synthesis

Cell-free protein synthesis under the direction of restriction fragments was accomplished as described previously (9, 10). The polypeptides synthesized were analyzed on 15% SDS-Tris-glycine polyacrylamide gels in the presence of 8M urea (10).

Nomenclature of promoter sites

Promoters are denoted by A or G depending on the ribonucleoside triphosphate with which RNA synthesis is initiated at the particular promoter site. The position of each promoter on the physical map is given in map units (one map unit corresponds to a genome length of 6400 bases) using the *Hind* II cleavage site as a reference (zero) point. Promoters have been numbered opposite to the direction of transcription.

RESULTS

1. Expression of gene III

In the absence of rho-factor only one strong termination site is operative on the M13 genome. This central termination site is located immediately distal to gene VIII but proximal to gene III (1, 3, 5, 21). Consequently, the expression of gene III may be accomplished either by a promoter in front of this gene or by a mechanism which allows leakage of RNA polymerase molecules through the termination site. Transcription studies have failed so far to detect a strong promoter in front of gene III, although the possibility that gene III is preceded by a promoter

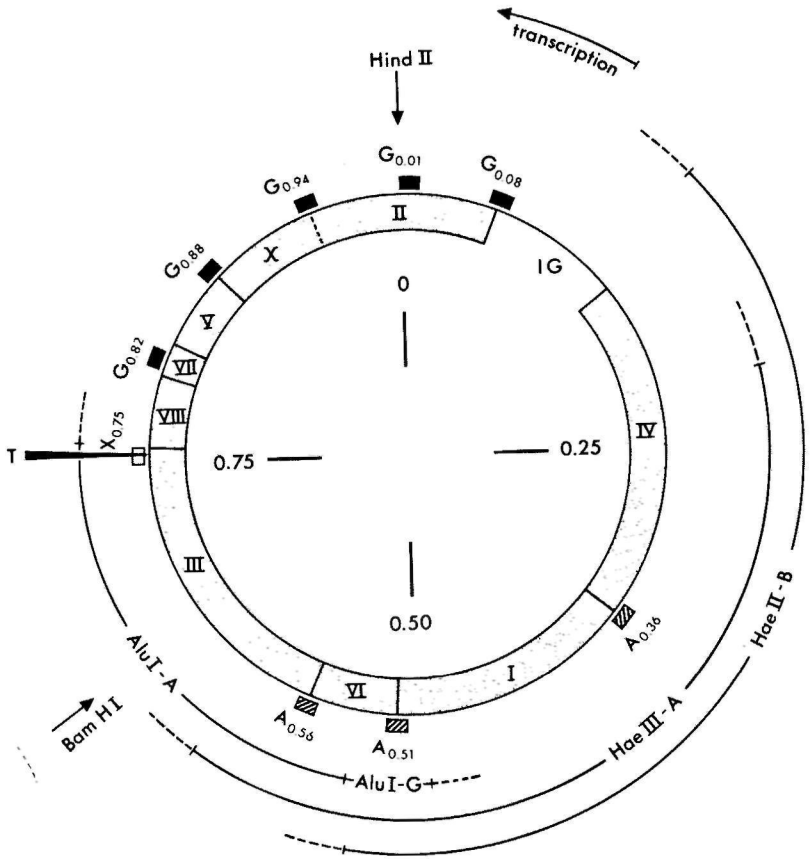


Fig. 1: Genetic map and physical maps of bacteriophage M13 DNA. The dotted circular area represents the genetic map. The uncompleted outer circles show the locations of restriction fragments containing specific parts of the M13 genome. The positions of the G-promoters are indicated with black bars, the positions of the A-promoters with hatched bars and the position of the gene III promoter with an open bar. The latter promoter forms an integral part of the central termination site of transcription (T). The direction of transcription is counterclockwise around the genetic map. The arrows indicate the single cleavage sites on M13 RF by endoR. Hind II and endoR. Bam H I, respectively. IG refers to the intergenic region between gene II and IV in which the replication origin for complementary and viral strand DNA synthesis is located.

† Note: The direction of transcription proceeds counterclockwise on the conventional genetic map. This is also the 5' to 3' polarity in the viral strand of each restriction fragment.

with a very low initiation capacity can not be excluded. Therefore we have studied that region of the genome which encompasses gene III in more detail.

In order to exclude the expression of gene III via a read-through mechanism, we have isolated a DNA template which codes for the complete gene III but contains none of the known G-promoter sites proximal to the central termination site. As shown in Fig. 1, the restriction fragment Alu I-A fulfills these requirements. From our transcription studies (3) and our nucleotide sequence data, presented elsewhere (6, Hulsebos and Schoenmakers, unpublished data), we know that the 5'-end of restriction fragment Alu I-A is located 46 nucleotides proximal to the central termination site. Consequently, this fragment does not contain any of the known G-promoters. The 3'-end of fragment Alu I-A is adjacent to restriction fragment Alu I-G which contains genetic markers for both genes VI and I (26). Since the order of genes is III-VI-I and fragment Alu I-G is relatively small, it is most likely that at the 3'-end of restriction fragment Alu I-A part of the coding information for the N-terminal part of gene VI is located.

As shown in Fig. 2a, transcription of fragment Alu I-A gives rise to the synthesis of two discrete RNA species. The largest of these RNA products has a length of about 1300 nucleotides (15S) whereas the smaller product is only about 200 nucleotides (6S) long. This observation already indicates that two RNA initiation sites are present on this fragment. Since fragment Alu I-A is 1430 base pairs long (26) and transcription proceeds in a counterclockwise direction on the physical map (12, 25), one may conclude that one RNA initiation site must be located near the 5' of this fragment. In a similar way one might conclude that the second RNA initiation site is located at a distance of about 200 nucleotides from the 3'-end of the DNA fragment, i.e. at 0.43 map units of the physical map. This observation is in complete agreement with our previous results from which we calculated that an A-promoter is located at 0.56 map units (4). Therefore, it is very probable that the small RNA product is initiated at this A-start promoter and is terminated at the 3'-end of fragment Alu I-A.

If the observed weak initiation of RNA synthesis at the 5'-end of fragment Alu I-A is due to a promoter which is located in front of gene III, fragment Alu I-A must be able to direct the synthesis of gene III protein in a DNA-dependent protein synthesizing system. To test this hypothesis, M13 RF was cleaved with an excess of restriction enzyme endoR. Alu I and the complete digest was directly translated in a coupled transcription-translation system. From the results which are presented in Fig. 3d, it is clear that the endoR. Alu I-cleaved RF is still able to direct the synthesis of a protein with a molecular weight identical to the molecular weight of gene III protein (cf. Fig. 3g). Since such a protein is not made when translation is carried out under the direction of a complete endoR. Alu I-digest of M13 RF carrying various am3-mutations (cf. Fig. 3a,b,c), it is almost certain that the high molecular weight protein is encoded by gene III. Translation of the isolated restriction fragment Alu I-A (Fig. 3e) confirms our presumption that this particular DNA fragment is responsible for the observed synthesis of gene III

Fig. 2: Autoradiogram of ^{32}P-labeled RNA species synthesized in vitro under the direction of restriction fragments of M13 RF. Products transcribed from (a) fragment Alu I-A; (b) from fragment Hae III-A. RNA synthesis was carried out in the presence of (α-^{32}P) CTP and was performed as described under Materials and Methods. The RNA products were analysed on vertical slab gels (1.9% acrylamide) in the presence of 7M urea. The nucleotide lengths of the RNA products indicated were estimated as described previously (4).

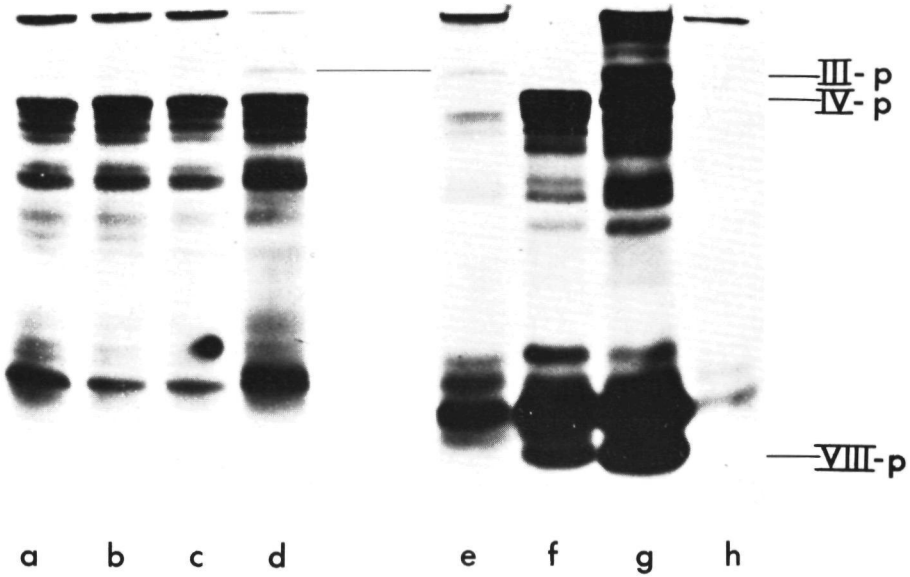


Fig. 3: Autoradiogram of (^{35}S) methionine-labeled polypeptides synthesized in vitro either under the direction of wild-type M13 RF, amber-mutant RF or under the direction of complete endoR. Alu I digests of M13 RF. Polypeptides synthesized under the direction of the complete endoR. Alu I-digests of M13 RF bearing (a) the am3-H5 mutation; (b) the am3-H4 mutation; (c) the am3-H1 mutation; (d) a complete endoR. Alu I digest of wild-type M13 RF; (e) purified fragment Alu I-A; (f) intact RF bearing the am3-H5 mutation; (g) intact wild-type RF; (h) products synthesized in the absence of exogeneous DNA. Protein synthesis in vitro was carried out as described (9, 10). An amount equivalent to 1 µg of replicative form DNA of each template was added per 25 µl of reaction mixture. The polypeptides synthesized in vitro were analyzed on 15% SDS-Tris-glycine gels (10). III-P, IV-P, etc. refer, respectively, to the positions of migration of the products of gene III, gene IV, etc. The mutants am3-H1 and am3-H4, and mutant am3-H5, have their amber mutation in different positions within gene III (26).

protein. Thus both transcription and translation studies strongly indicate the existence of a (weak) promoter in front of gene III.

Three of the promoters which are located on the M13 genome initiate RNA synthesis which start with pppA. Although the map positions of all these A-promoters are known exactly, up to now only the strong promoter $A_{0.36}$ has been localized on the genetic map of the M13 genome. Its position was found immediately in front of gene IV (10, 25).

In order to determine also the positions of the remaining A-promoters, we have studied the transcription properties of restriction fragment Hae III-A. This fragment encompasses the C-terminal part of gene III, the entire gene VI and gene I and about 1000 base pairs of the N-terminal part of gene IV (24, 25). Upon transcription of this DNA fragment, three RNA products are formed the sizes of which are about 900, 1700 and 2100 nucleotides, respectively (Fig. 2b). Given the assumption that termination of transcription occurs at the 5'-end of the codogenic strand of each fragment, the estimated lengths of these RNA species confirm a starting point of RNA chain-growth at the previously identified promoters $A_{0.36}$, $A_{0.51}$ and $A_{0.56}$ (4). Since the latter two promoters are located on the genetic map in a region where also several amber mutations of gene VI and gene I have been mapped (24, 26), it is tempting to speculate that these two A-promoters are located immediately in front of the genes VI and I. If true, restriction fragment Hae III-A must be able to direct the synthesis of both proteins. Translation of fragment Hae III-A in a coupled transcription-translation system clearly shows the presence of a protein which comigrates with gene I protein synthesized under the direction of intact M13 RF (Fig. 4b and f). The synthesis of this protein, i.e. gene I protein, also occurs with M13 RF that has been cleaved once with endoR. Hind II (Fig. 4d). Synthesis of gene I protein, however, is not directed by M13 RF derived from M13 phage DNA carrying an amber mutation in gene I (Fig. 4c). We are confident, therefore, that fragment Hae III-A is capable of directing the synthesis of gene I protein and that at least one of the two A-promoters must be located in front of gene I. The presence of gene VI protein among the translational products, however, could not be demonstrated. Its synthesis is

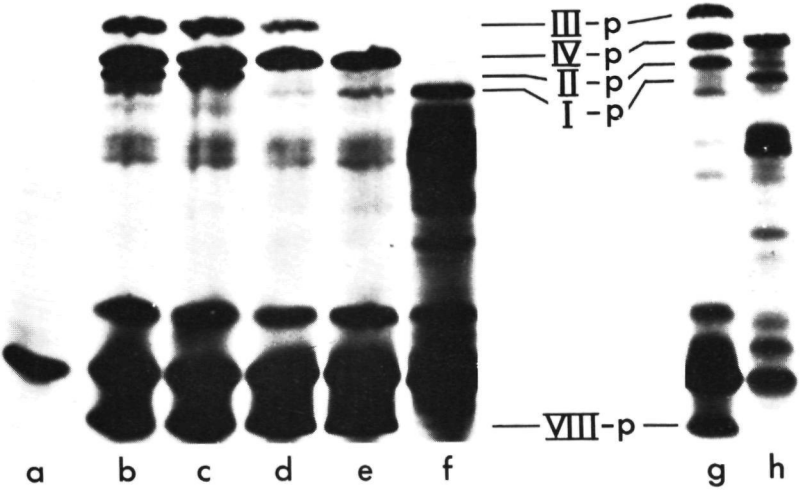


Fig. 4: Autoradiogram of (^{35}S) methionine-labeled polypeptides synthesized *in vitro* under the direction of either wild-type M13 RF, amber-mutant RF or various restriction fragments. (a) Products synthesized in the absence of exogeneous DNA; (b) products synthesized under the direction of intact wild-type M13 RF I; (c) M13 RF I bearing the *am*1-H6 mutation; (d) M13 RF cleaved with endoR. *Hind* II; (e) M13 RF cleaved with endoR. *Bam*H 1; (f) purified fragment *Hae* III-A; (g) M13 RF bearing the *am*1-H3 mutation; (h) purified fragment *Hae* II-B. Protein synthesis was carried out as described previously (10). Further conditions are described in the legend of Fig. 3. I-P, II-P, etc. refer, respectively, to the positions of migration of the products of gene I, gene II, etc.

neither apparent on fragment *Hae* III-A nor on intact M13 RF. In fact this is not surprising since the synthesis of this protein has not been detected either *in vitro* (10, 12) or *in vivo* (8, 20). Due to this failure, a conclusion whether gene VI is also equipped with a promoter can not be drawn.

To establish whether gene I is preceded by promoter $A_{0.56}$ or by promoter $A_{0.51}$ translation experiments were performed with restriction fragment *Hae* II-B. From its position on the genetic and physical map (26), we know that this fragment contains the complete gene I and gene IV (Fig. 1). Nucleotide sequence analyses have shown that this fragment contains in addition the coding

information for 49 amino acid residues of the C-terminal end of gene VI protein (Wezenbeek and Schoenmakers, unpublished results). Hence promoter $A_{0.56}$ is not encompassed in fragment Hae II-B. As shown in Fig. 4h, upon translation of fragment Hae II-B not only the complete gene IV protein is synthesized but also the synthesis is apparent of a protein which comigrates with gene I protein synthesized under the direction of intact M13 RF (cf. Fig. 4b). The latter protein is absent upon translation of M13 RF carrying an amber mutation in gene I (Fig. 4g). From these results we conclude that the capacity of directing the synthesis of gene I protein is still preserved on fragment Hae II-B and consequently promoter $A_{0.51}$ is the one which must be located immediately in front of gene I.

Due to the fact that promoter $A_{0.56}$ is positioned upstream of promoter $A_{0.51}$ and gene VI is located proximal to gene I (Fig. 1) it is attractive to postulate that $A_{0.56}$ is the promoter which is positioned in front of gene VI. However, due to the absence of detectable synthesis of gene VI protein in the DNA-dependent cell-free system no direct evidence can be provided for this assumption. On the other hand, the distance between promoter $A_{0.56}$ and promoter $A_{0.51}$ is about 350 base pairs (4). From our nucleotide sequence studies we know now that gene VI encompasses only 336 base pairs, and that the regions which code for gene VI ard gene I proteins are only separated by a single base pair (Wezenbeek and Schoenmakers, unpublished results). Hence we feel justified in concluding that promoter $A_{0.56}$ is located in front of gene VI. The synthesis of an RNA species of only 200 nucleotides on fragment Alu I-A (Fig. 2a), which contains only the N-terminal part of gene VI, strongly supports this conclusion.

Based upon genetic complementation studies it has been suggested that the M13 genes III, VI and I form a regulatory unit which only can be expressed via a single polycistronic mRNA (11, 15). To test whether this is true, use was made of M13 RF which has been cleaved with endoR. Bam H I. The latter enzyme has only a single recognition site on the M13 RF molecule and cleaves the DNA genome within gene III (Fig. 1; Wezenbeek and Schoenmakers, unpublished results). Translation of endoR. Bam H I-cleaved RF in a DNA-dependent protein synthesizing system shows,

as expected, that the synthesis of gene III protein is absent (Fig. 4e). If the expression of genes III, VI and I occurs on polycistronic mRNAs, as suggested by the polarity effects, synthesis of gene I protein should also be abolished on this template. However, as is shown in Fig. 4e, this is not the case. From these results we conclude that initiation of transcription of genes VI and I at the promoter sites immediately preceding these genes is retained and that transcription of gene III is not a prerequisite for the expression of the distal genes.

DISCUSSION

In previous communications we have used transcription and translation studies of restriction fragments to demonstrate the existence of promoters in front of particular genes. The capability of a DNA fragment or of its transcript to direct the *in vitro* synthesis of a phage-specific polypeptide was used as an indication of whether in front of the gene coding for this polypeptide a promoter is located. Using this approach, we could demonstrate that the promoters $G_{0.82}$, $G_{0.88}$, $G_{0.08}$ and $A_{0.36}$ are, respectively located immediately in front of genes VIII, V, II and IV. In addition it was found that the promoters $G_{0.94}$ and $G_{0.01}$ are intragenic RNA initiation sites which are located within gene II. For promoter $G_{0.94}$ it furthermore was found that this is located in front of a gene, i.e. gene X whose nucleotide sequence completely overlaps with the C-terminal end of gene II.

Following the same strategy we now have presented evidence that the promoters $A_{0.56}$ and $A_{0.51}$ belong to the genes VI and I and that there also exists a promoter, i.e. $X_{0.75}$, immediately in front of gene III.

Although our protein synthesis data clearly demonstrated the existence of a gene III promoter, the inaccuracy of length measurements of RNA chains on polyacrylamide gels did not allow a more precise location of this promoter to be made. Theoretically, the gene III promoter can be located either immediately distal or proximal to the central termination site. In the latter case the terminator might also function as an attenuator which regulates the expression of gene III. A proximal position, however,

is very unlikely since in that case the gene III promoter should generate very short RNA chains which are prematurely terminated at the termination site. Upon transcription of the appropriate restriction fragment such short RNA chains with a characteristic U_8-track at their 3'-OH terminus (cf. 21) have never been detected among the transcriptional products (Edens, unpublished data). Recent nucleotide sequence analysis of Takanami and co-workers (personal communication) have now revealed that the gene III promoter of phage fd forms an integral part of the central termination site. Since the nucleotide sequence of the terminator region on the M13 genome has been found to be completely identical to the one deduced for phage fd (6, 21), it is almost certain that the observed gene III promoter on the M13 genome is also located within the central terminator region. Such a position also explains the observed diminished synthesis of gene III protein on endoR. Alu I-treated RF as compared to intact superhelical RF (cf. Fig. 3). The endoR. Alu I cut is 46 nucleotides proximal to the central terminator and, hence, might influence the gene III promoter function. From our nucleotide sequence analyses we learned that the primary 'binding sequence' (cf. 16, 19) of the gene III promoter is still preserved in fragment Alu I-A, but that the second, so called 'recognition site' (16), although not being an absolute necessity for proper RNA polymerase binding and RNA chain initiation (14), is missing. The diminished gene III protein synthesis is therefore most probable caused by a partly destroyed promoter function.

Additional effects on gene III protein synthesis can not be ruled out however. Previously we demonstrated that termination of transcription at the central termination site - at least under the in vitro conditions applied - is not stringent (4, 5). Since the endoR. Alu I cut is immediately proximal to the central terminator this cut will additionally exert a reduction of gene III protein synthesis due to a switch-off of read-through of transcription at this particular site. To what extent this read-through contributes to the expression of gene III can not be answered yet since proper (linear) RF DNA templates can not yet be constructed on which read-through is completely eliminated without concomitant disturbance of gene III promoter function. A more general

problem underlying such *in vitro* transcription-translation studies is that RNA initiation frequencies will differ on DNA templates with different configurations. It has already been demonstrated by Seeburg *et al.*, (19) that the RNA polymerase binding capacity is diminished at least tenfold from supercoiled to relaxed RF. In our opinion, this reduction in complex formation and hence in RNA-initiation is not an overall effect but might be completely different for each individual promoter. This is clearly recognized by comparing the patterns obtained with intact RF (Fig. 4b) with those obtained from endoR. Hind II - treated RF (Fig. 4d) and endoR. BamH I-treated RF (fig. 4e). The amount of gene II protein is high on supercoiled RF and is absent, as expected, on RF in which an endoR. Hind II cut is introduced into gene II. Introduction of an enzyme cut into a completely different region, i.e. within gene III, diminishes also the synthesizing capacity of gene II protein to a very large extent. Therefore, a direct comparison of the initiation frequency for a given promoter on various linear and supercoiled RF molecules can not be made and limits the interpretations of results in which these different DNA templates have been applied.

Another aspect of our results concerns the observed polarity among the genes III, VI and I. The operon model, which has been postulated to be pertinent for these genes (11,15), implies that the introduction of a cleavage whithin gene III also prevents the expression of the distal genes VI and I. Our results of coupled transcription and translation, however, suggest the opposite since an interruption of gene III transcription does not abolish the synthesis of gene I protein. A result which is in accordance with the existence of promoters in front of genes VI and I. If these promoters also function as such in the infected cell the operon model is no longer applicable for these genes and other regulatory mechanisms have to be reconsidered. Further studies like *in vitro* rearrangements of genes and a detailed analysis of the *in vivo* RNA species are needed to clarify these contradictory results on polarity.

REFERENCES

1. Chan, T., Model, P. and Zinder, N.D. (1975). In vitro protein synthesis directed by separated transcripts of bacteriophage f1 DNA. *J. Mol. Biol.*, 99, 369-382.
2. Denhardt, D.T. (1975). The single-stranded DNA phages. *CRC Critical Reviews in Microbiology*, pp. 116-123, The Chemical Rubber Company, Cleveland, Ohio.
3. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Physical mapping of the central terminator for transcription on the bacteriophage M13 genome. *Nucleic Acid Res.*, 2, 1811-1820.
4. Edens, L., van Wezenbeek, P., Konings, R.N.H. and Schoenmakers, J.G.G. (1976). Mapping of promoter sites on the genome of bacteriophage M13. *Eur. J. Biochem.*, 70, 577-587.
5. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). A cascade mechanism of transcription in bacteriophage M13 DNA. *Virology*, 86, 354-367.
6. Edens, L. (1978). Regulation of expression of the bacteriophage M13 genome. Initiation and termination of transcription. *Ph.D. Thesis*, University of Nijmegen.
7. Giacomani, P.U., Delain, E. and le Pecq, J.G. (1977). Electron microscopy analysis of the interaction between Escherichia coli DNA-dependent RNA polymerase and the replicative form of phage fd DNA. I. Mapping of the binding sites. *Eur. J. Biochem.*, 78, 205-213.
8. Henry, T.J. and Pratt, D. (1969). The proteins of bacteriophage M13. *Proc. Nat. Acad. Sci. U.S.A.*, 62, 800-807.
9. Konings, R.N.H., Ward, R., Francke, B. and Hofschneider, P.H. (1970). Gene order of RNA bacteriophage M12. *Nature*, 266, 604-607.
10. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975). Identification and characterization of the in vitro synthesized products of the bacteriophage M13. *J. Virol.*, 15, 570-584.
11. Lyons, L.B. and Zinder, N.D. (1972). The genetic map of the filamentous bacteriophage f1. *Virology*, 49, 45-60.

12. Model, P. and Zinder, N.D. (1974). In vitro-synthesis of bacteriophage f1 proteins. *J. Mol. Biol.*, 83, 231-251.
13. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, M. (1975). Studies on bacteriophage fd DNA. II. Localisation of RNA-initiation sites on the cleavage map of the fd genome. *J. Mol. Biol.*, 95, 33-44.
14. Okamoto, T., Sugimoto, K., Sugisaki, H. and Takanami, M. (1977). DNA regions essential for the function of bacteriophage fd promoter. *Nucleic Acid Res.*, 4, 2213-2222.
15. Pratt, D., Tzagoloff, H., Erdahl, W.S. and Henry, T.J. (1967). Conditional lethal mutants of coliphage M13. In: *Molecular Biology of Viruses.* (J.S. Colter and W. Parachych, eds.), pp. 219, Academic Press, New York.
16. Pribnow, D. (1975). Bacteriophage T7 early promoters: nucleotide sequences of two RNA polymerase binding sites. *J. Mol. Biol.*, 99, 419-443.
17. Ray, D.S. (1977). Replication of filamentous bacteriophages. *Comprehensive Virology.* (H. Fraenkel-Conrat and R.R. Wagner, eds.), pp. 105-178, Plenum Publishing Corporation, New York.
18. Seeburg, P.H. and Schaller, H. (1975). Mapping and characterization of promoters in phage fd, f1 and M13. *J. Mol. Biol.*, 92, 261-277.
19. Seeburg, P.H., Nüsslein, C. and Schaller, H. (1977). Interaction of RNA polymerase with promoters from bacteriophage fd. *Eur. J. Biochem.*, 74, 107-113.
20. Smits, M.A., Simons, G., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). Expression of bacteriophage M13 DNA in vivo. I. Synthesis of phage-specific RNA and proteins in minicells. *Biochem. Biophys. Acta,* XX-000-000.
21. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1977). Studies on bacteriophage fd DNA. IV. The sequence of messenger RNA for the Major Coat Protein gene. *J. Mol. Biol.*, 110, 487-507.
22. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1975). Studies on bacteriophage M13 DNA. I. A cleavage map of the M13 genome. *Eur. J. Biochem.*, 53, 547-558.

23. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1976). Cleavage maps of the filamentous bacteriophages M13, f1, fd and ZJ/2. *J. Virol.*, 18, 1024-1039.

24. Van den Hondel, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Studies on bacteriophage M13 DNA. II. The gene order of the M13 genome. *Eur. J. Biochem.*, 53, 559-567.

25. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). Regulation of gene activity in bacteriophage M13 DNA: coupled transcription and translation of purified genes and gene fragments. *Virology*, 67, 487-497.

26. Van den Hondel, C.A., Pennings, L. and Schoenmakers, J.G.G. (1976). Restriction enzyme cleavage maps of bacteriophage M13: existence of an intergenic region. *Eur. J. Biochem.*, 68, 55-70.

Note: This article has been accepted by the *Journal of Virology* in a form in which the nomenclature of the promoter sites has been changed from that used in the above text in accordance with recently accepted international rules.

CHAPTER VIII

EFFECT OF TERMINATION FACTOR RHO ON THE *IN VITRO* TRANSCRIPTION OF THE BACTERIOPHAGE M13 GENOME

EFFECT OF TERMINATION FACTOR RHO ON THE IN VITRO TRANSCRIPTION OF THE BACTERIOPHAGE M13 GENOME

ABSTRACT

Transcription studies were performed in the presence and absence of rho protein to examine the effect of this termination factor on the in vitro transcription of the M13 RF molecule. By using intact RF, as well as several restriction fragments, as templates in the transcription reaction, it was demonstrated that in addition of moderate amounts of rho had only a minor effect on the termination efficiency of the central termination signal as well as on the synthesis of the various pppG-initiated RNA chains. However, the addition of rho was found to seriously influence the sizes of all pppA-initiated RNA chains. From these observations, and from the length of the rho-terminated, pppA-initiated RNA products, it is concluded that an efficient, rho-dependent termination signal(s) is present in the border of gene IV and the intergenic region.

INTRODUCTION

In vitro transcription studies on replicative form DNA (RF-I) of the filamentous phages M13, fd and f1, have indicated that these templates are transcribed according to a multi-promoter single-terminator mechanism (for a recent review see 1). Transcription of the circular DNA molecule with E.coli RNA polymerase holoenzyme in the absence of transcription termination factor rho thus results in a set of discrete RNA species which are initiated at the various promoters, but are all terminated at one unique site, the so called 'central' termination site. All the RNA species generated, therefore, share the coding information for that part of the genome which is located immediately proximal to this central termination signal. Five of the RNA species generated range in size from 8S upto 19S and are initiated at the G-promoters $G_{0.82}$, $G_{0.88}$, $G_{0.94}$, $G_{0.01}$ and $G_{0.08}$ respectively. Three larger RNA species, with sizes of 23S and 26S, are initiated at the A-promoters $A_{0.36}$, $A_{0.51}$ and $A_{0.56}$ (2). The largest RNA species formed on the RF-I molecule is initiated at a promoter ($X_{0.75}$) which partially overlaps the central termination region (3, 4). The length of the latter RNA species (30S) represents, therefore, one round of transcription.

In contrast to our knowledge of the transcription mechanism operative *in vitro*, there are only a few experimental data available concerning the mechanism by which the RF molecule is transcribed *in vivo*. The hybridisation studies described by Jacob *et al.*, (5) suggest that the maximum size of phage M13-specific RNA species isolated from infected *E.coli* cells is considerably smaller than the maximum size which has been observed for RNA products synthesized *in vitro*. Additionally, analysis of phage-specific RNA species made in RF harbouring *E.coli* mini-cells has demonstrated that among these products there are no RNA species detectable which are larger than 20S (i.e. about 2300 nucleotides; 6). Both observations suggest, therefore, that there exist termination signal(s) on the M13 genome which have not yet been detected by *in vitro* studies. Since to date the *in vitro* transcription studies on M13 DNA have been carried out in the absence of termination factor rho, we now describe the effect of rho on the *in vitro* transcription process. The results of these studies suggest that there are two transcription termination signals present on the M13 genome. One of these signals is rho-independent and identical to the previously located central termination site. The other is only operative in the presence of rho and, due to its position on the phage genome, this signal affects the sizes of only the larger RNA species.

MATERIALS AND METHODS

Materials

Escherichia coli RNA polymerase holoenzyme was a generous gift from Dr. R. van Meteren (University of Leiden). Lambda cb_2 DNA was kindly made available by Dr. H. Pannekoek (University of Leiden). The restriction endonucleases from *Haemophilus influenza Rd* (endo R. *Hind* II), *Haemophilus aphirophilus* (endo R. *Hap* II), *Haemophilus aegyptius* (endo R. *Hae* III) were isolated as described previously (7, 8). The restriction endonucleases from *Bacillus amyloliquefaciens* (endo R. *Bam*. H I) was purchased from New England Biolabs, U.S.A.

Replicative form I DNA and restriction fragments

The method for the preparative isolation and purification of M13 replicative form I DNA has been described (9). Digestion of it with restriction endonucleases has been described elsewhere (8). After digestion, the fragments were separated by electrophoresis on 3.0% discontinuous polyacrylamide slab gels and further purified as described by Van den Hondel et al., (10).

Purification of rho

Rho was purified from E.coli MRE 600 by a combination of procedures described by Roberts (11) and Darlix (12). Homogenisation of cells and phosphocellulose chromatography was carried out according to Darlix whilst DEAE cellulose chromatography and subsequent glycerol gradient centrifugation were performed as described by Roberts. Upon electrophoresis through 13% SDS-Tris-Glycine polyacrylamide gels in the presence of 8M urea (13) the isolated rho protein gave a single band corresponding with a molecular weight of about 50.000.

RNA synthesis in vitro

RNA synthesis in vitro was performed in a standard transcription reaction mixture (0.1 ml) containing: 4 µmol of Tris-HCl (pH 7.9), 5 µmol of KCl, 0.8 µmol of $MgCl_2$, 0.1 µmol of dithiotreitol, 0.01 µmol of EDTA, 0.1% Tween-80 and 25 µg gelatin. This reaction mixture was supplemented with 0.2 pmol of M13 DNA restriction fragment, 2.0 pmol of E.coli RNA polymerase holoenzyme and, if transcription was carried out in the presence of rho, 20 pmol of the isolated rho protein. After a pre-incubation period of 5 min at 37°C, ribonucleoside triphosphates were added to a final concentration of 400 µM, except for the (α^{32}P)-labeled CTP whose concentration was 40 µM. After incubation for 5 min at 37°C, the CTP concentration was raised to 400 µM. After another 5 min the reactions were terminated by the addition of 10 µl of 10% SDS and 0.1 ml of phenol saturated with a buffer containing 0.01 M Tris-HCl (pH 7.6) and 0.001 M EDTA. Carrier

RNA was added (10 µg) and the reaction mixtures were extracted with phenol. Subsequently the RNA was precipitated twice with 2.5 volumes of cold ethanol and finally dissolved in 20 µl of deionized formamide.

Transcription termination assay

The termination activity of the isolated rho protein was determined in an *in vitro* transcription assay in which lambda *cb2* DNA was used as a template. The assay procedure was carried out in the standard transcription reaction mixture supplemented with 0.07 pmol of template DNA, 2.0 pmol of *E.coli* RNA polymerase holoenzyme and a variable amount of rho. The synthesis of RNA and the termination of the transcription reaction were carried out essentially as described under RNA synthesis *in vitro*. The RNA products synthesized under the different conditions were analysed by gel electrophoresis on 3% acrylamide gels in the presence of 7 M urea (14)and the amount of 4S RNA synthesized was determined. When less than 5 pmol of the isolated rho protein were added to the reaction mixture, the synthesis of 4S RNA was hardly detectable. The addition of larger amounts of rho resulted,however, in a considerable stimulation of the synthesis of this rho-terminated RNA product (11).

Polyacrylamide gel electrophoresis

The *in vitro* synthesized RNA species were analysed on either 2.0% polyacrylamide slab gels in the presence of 7M urea (14) or on 1.5% agarose gels in the presence of 6M urea. After electrophoresis, the wet slab gels were exposed to X-ray film (Kodak XRI).

Nomenclature of promoter sites

Promoters are denoted by A or G depending on the ribonucleoside triphosphate with which RNA synthesis is initiated at the particular promoter site. The position of each promoter on the physical map is given in map units (one map unit corresponds to

a genome length of 6400 bases) using the Hind II cleavage site as a reference (zero) point. Promoters have been numbered opposite the direction of transcription.

RESULTS

Effect of rho on the transcription of intact M13 RF

Transcription of intact M13 RF with E.coli RNA polymerase holoenzyme results in the synthesis of several discrete RNA products. Analysis of these products by electrophoresis on either 2.0% acrylamide gels (Fig. 2c) or 1.5% agarose gels (Fig. 2b), clearly demonstrates that, among the RNA species synthesized, the species initiated at the promoters $G_{0.82}$ (8S; 370 nucleotides), $G_{0.94}$ (14S; 1200 nucleotides), $G_{0.08}$ (19S; 2000 nucleotides) and $A_{0.36}$ (23S; 3900 nucleotides) are predominant (Fig. 1). However, as shown in Figs. 2d, e, f, addition of purified termination factor rho to the transcription system significantly affects this transcription pattern. Upon increasing the rho concentration, the synthesis of several new RNA species becomes apparent whereas the synthesis of the larger RNA species is reduced. Under the low salt conditions used (50 mM KCl), an optimum effect is reached after the addition of about 20 pmol rho. This amount of rho has, therefore, been used in all transcription experiments. Under these standard conditions the synthesis of pppG-initiated RNA species (8S up to 19S) remains unaffected whereas the synthesis of pppA-initiated RNA species (23S up to 26S) is strongly reduced (Figs. 2a,f). In addition, rho induces the synthesis of several new RNA species. On the basis of these two observations, we infer that rho-dependent termination signals are operating on the genome of bacteriophage M13. In order to map these signals, transcription studies were performed in the presence and absence of rho on restriction fragments which were derived by digestion of the RF-I molecule with endo R. Hind II, endo R. Bam H. I, endo R. Hae III and endo R. Hap II.

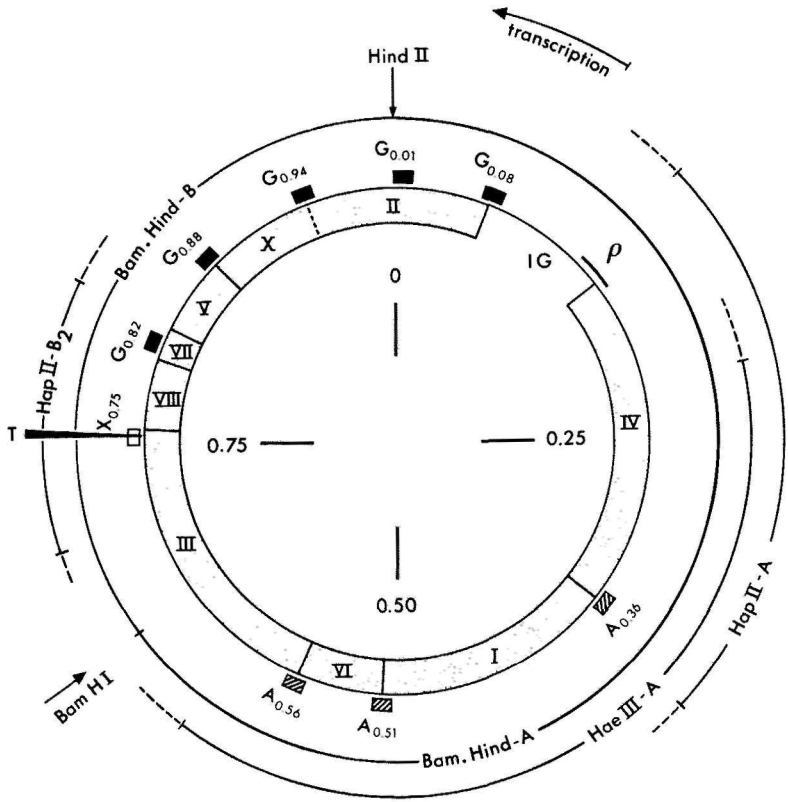


Fig. 1: Genetic map and physical map of bacteriophage M13 DNA. The dotted circular area represents the genetic map. The uncompleted outer circles show the locations of restriction fragments containing specific parts of the M13 genome. The positions of the G-promoters are indicated with black bars, the positions of the A-promoters with hatched bars, and the position of the gene III-promoter with an open bar. The latter promoter partially overlaps the central termination site of transcription ($T_{0.75}$). The direction of transcription is counterclockwise around the genetic map. The arrows indicate the positions of the single cleavage sites by endo R. Hind II and endo R. Bam H I, respectively. IG refers to the intergenic region between gene II and IV in which the replication origin for parental RF and viral DNA synthesis is located. The DNA region where efficient, rho-dependent, termination signals are located is indicated by ρ.

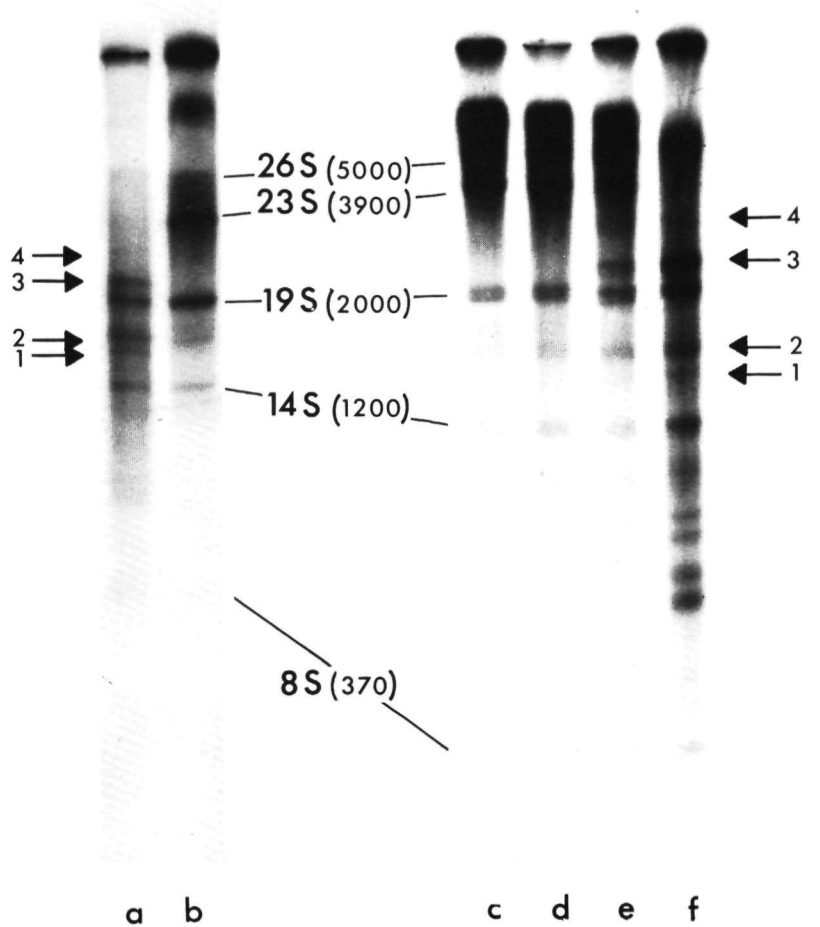


Fig. 2: Autoradiogram of (α^{32}P)-labeled species synthesized in vitro under the direction of phage M13 replicative form DNA in either the presence or absence of rho. The RNA products were synthesized (a) in the presence of 20 pmol rho; (b) in the absence of rho; (c) in the absence of rho; (d) in the presence of 5 pmol rho; (e) in the presence of 15 pmol rho; (f) in the presence of 30 pmol rho. RNA synthesis was carried out in the presence of (α^{32}P)-labeled CTP and was performed as described under Materials and Methods. The RNA products were analysed on vertical slab gels of either 1.5% agarose in the presence of 6 M urea (lanes a and b) or on 2.0% acrylamide in the presence of 7M urea (lanes c-f). The approximate sizes of the RNA products (number of nucleotides) are indicated in parenthesis and were estimated as described previously (5). The RNA species numbered 1 through 4 refer to rho-terminated products (see text).

Effect of rho on the transcription pattern of restriction fragment Bam. Hind-A

Both restriction endonucleases Hind II and Bam H I have only a single cleavage site on the M13 genome. The double-stranded cut introduced by endo R. Hind II is located within gene II at a position which has been defined as the zero point of the physical map. (Fig. 1; 8). The cleavage site of endo R. Bam H I is located within gene III, about 700 nucleotides distal of the central termination site (Van Wezenbeek and Schoenmakers, unpublished results). Thus, upon cleavage of M13 RF-I with both restriction enzymes, two DNA fragments result which are designated Bam. Hind-A and Bam. Hind-B respectively. Of these two fragments the larger, i.e. Bam. Hind-A, contains gene VI, gene I and gene IV whereas gene III and gene II are encompassed only partially on this fragment (Fig. 1).

Since the cleavage site of endo R. Hind II is located very close to promoter $G_{0.01}$ (15, 2), transcription of fragment Bam. Hind-A in the absence of the rho results in the synthesis of only four discrete RNA species (Fig. 3f). The synthesis of all of these RNA species is presumed to be terminated at the 3'-end † of the restriction fragment (2, 16). The smallest RNA species formed on fragment Bam. Hind-A (Δ19S) is initiated at the G-promoter $G_{0.08}$ in front of gene II and has a length of about 450 nucleotides. The three remaining RNA species are initiated at A-promoters which are located in front of gene IV, gene I and gene VI respectively. Of these transcripts the major one (Δ23S; 2000 nucleotides) is initiated at the promoter $A_{0.36}$ whereas the other two (Δ26S; about 3200 nucleotides) are initiated at the promoters $A_{0.51}$ and $A_{0.56}$ respectively (2, 4).

When transcription of fragment Bam. Hind-A is carried out in the presence of rho, the RNA product initiated at promoter $G_{0.88}$ (Δ19S) remains unaffected (Fig. 3e). The synthesis of the three other RNA species, however, is strongly reduced and instead

† Note: The direction of transcription proceeds counterclockwise on the conventional genetic map. This is also the 5' to 3' polarity in the viral strand of each restriction fragment.

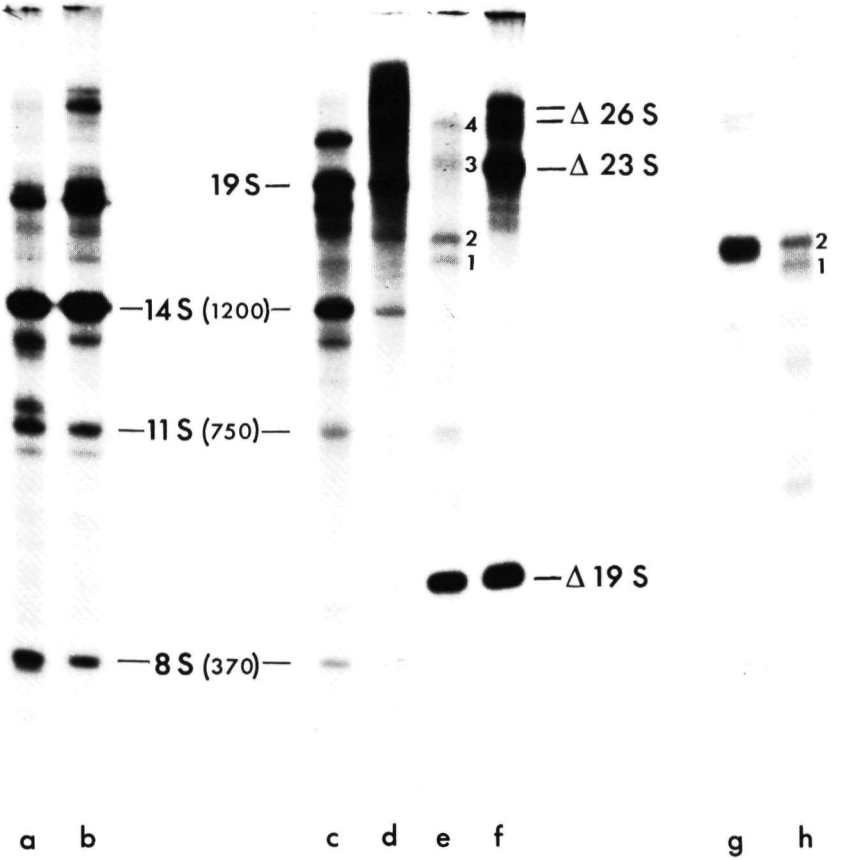


Fig. 3: Autoradiogram of RNA species synthesized in vitro under the direction of various restriction fragments and in either the presence or absence of rho. The RNA products were synthesized under the direction of (a) fragment Bam. Hind-B in the presence of 20 pmol rho; (b) fragment Bam. Hind-B in the absence of rho; (c) M13 RF cleaved with endo R. Bam H I in the absence of rho; (d) intact M13 RF in the absence of rho; (e) fragment Bam. Hind-A in the presence of 20 pmol rho; (f) fragment Bam. Hind-A in the absence of rho; (g) fragment Hap II-A in the absence of rho; (h) fragment Hap II-A in the presence of 20 pmol rho. RNA synthesis was carried out as described under Materials and Methods except for those experiments in which intact M13 RF (lane d) and M13 RF cleaved with endo R. Bam H I (lane c) have been used as templates. In the latter two experiments RNA synthesis was carried out in the presence of 15 µmol of KCl. The RNA products were analysed on vertical slab gels of 2.0% acrylamide in the presence of 7M urea. The approximate sizes of the RNA products are given in

nucleotides and indicated in parenthesis. The prefix Δ indicates that the RNA products synthesized are prematurely terminated at the endo R. Hind II cleavage site. The RNA species numbered 1 through 4 refer to rho-terminated RNA products (see text).

several smaller RNA species are now observed. These results not only confirm the experimental data obtained from the transcription of RF-I (cf. Fig. 2) but also indicate the presence of rho-dependent termination signals on fragment Bam. Hind-A. Since the RNA species initiated at promoter $G_{0.08}$ is the only transcript the size of which is not affected by the addition of rho, this observation suggests furthermore that these termination signals are located between promoter $G_{0.08}$ and promoter $A_{0.36}$ (Fig. 1).

Comparison of Fig. 3e with Fig. 3f shows that the synthesis of the newly induced RNA species is diminished. This is in contrast to the Δ19S RNA species, the synthesis of which is not affected significantly by the addition of rho. Although the reason for the reduced synthesis of the rho-terminated, pppA-initiated RNA species is not known, there is evidence to suggest that under in vitro conditions rho stimulates minor termination events at many sites, thereby reducing the amount of the predicted, rho-terminated RNA product(17; Fig. 2f, 3e).

Effect of rho on the transcription pattern of restriction fragment Bam. Hind-B

Restriction fragment Bam. Hind-B contains the promoters $G_{0.94}$, $G_{0.88}$, $G_{0.82}$ and $X_{0.75}$ which precede gene X, gene V, gene VIII and gene III respectively. (Fig. 1; 18, 4, 3). This restriction fragment also contains the central termination site ($T_{0.75}$) which is located immediately distal of gene VIII (19, 16).

Upon transcription of fragment Bam. Hind-B in the absence of rho, three pppG-initiated RNA species are formed all of which are terminated immediately distal to gene VIII(18, Fig. 3b).The sizes of these three RNA species range from 8S (370 nucleotides) up to 14S (1200 nucleotides). Fragment Bam. Hind-B does, however, direct the synthesis of some additional RNA species which are also apparent among the RNA products synthesized with endo

R. Bam. Hl—treated RF (Fig. 3c). Previously we have argued that these additional RNA species represent products which are the result of leakage of RNA polymerase molecules through the central termination site (18).

Upon transcription fragment Bam. Hind-B in the presence of rho, the synthesis of the 8S RNA, 11S RNA and 14S RNA species is clearly recognized (Fig. 3a). A similar set of RNA species was obtained by transcription of intact RF in the presence of rho (cf. Fig. 2) indicating that the presence of a strong, rho-dependent, termination signal in that part of the M13 genome which contains the genes X, V, VII and VIII is unlikely. However, we have noted that upon transcription of restriction fragments covering this particular part of the M13 genome, rho reproducably induces the synthesis of some minor RNA species. Whether these minor products are the result of unspecific, rho-mediated, transcription events or originate from defined but weak, rho-dependent, termination signals cannot be ascertained yet.

Effect of rho on the transcription patterns of the restriction fragments Hae III-A and Hap II-A

Our transcription studies with intact RF-I and with the restriction fragments Bam. Hind-A and Bam. Hind-B indicated that rho strongly affects the pppA-initiated RNA chains whereas the pppG-initiated molecules remain unaltered. This observation implies the presence of strong, rho-dependent, termination signal(s) in that region of the M13 genome which is located between promoter $G_{0.08}$ and promoter $A_{0.36}$. In order to locate these rho-dependent termination site(s) more precisely, use was made of the restriction fragments Hae III-A and Hap II-A (Fig. 1).

Previously we have shown that restriction fragment Hae III-A contains the A-promoters $A_{0.56}$, $A_{0.51}$ and $A_{0.36}$ which are located in front of gene VI, gene I and gene IV respectively (4). Consequently, transcription of this restriction fragment in the absence of rho results in the synthesis of three RNA species (Fig. 4b). The smallest one of the generated RNA species has a length of about 900 nucleotides and encodes the N-terminal part of gene IV. The other two RNA species are about 2100 and 1750

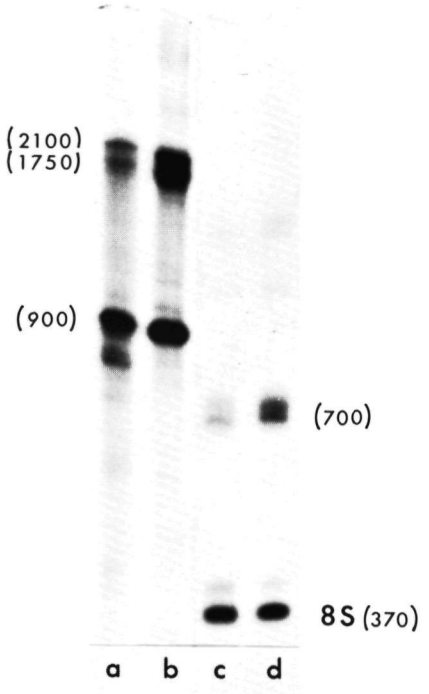


Fig. 4: Autoradiogram of RNA species synthesized *in vitro* under the direction of various restriction fragments in either the presence or absence of rho. The RNA products were synthesized under the direction of (a) fragment *Hae* III-A in the presence of 20 pmol rho; (b) fragment *Hae* III-A in the absence of rho: (c) fragment *Hap* II-B_2 in the presence of 20 pmol rho; (d) fragment *Hap* II-B_2 in the absence of rho. RNA synthesis was carried out as described under Materials and Methods and the products synthesized were analysed by electrophoresis through 2.0% acrylamide gels under denaturing conditions. The approximate lengths of the synthesized products are given in nucleotides and indicated in parenthesis.

nucleotides long and are initiated at the weak promoters $A_{0.56}$ and $A_{0.51}$ respectively. If restriction fragment *Hae* III-A contains a rho-dependent termination signal and if this signal is positioned downstream of promoter $A_{0.36}$, then it is expected that upon transcription of this fragment in the presence of rho, all three pppA-initiated RNA species are reduced in length. However, as shown in Fig. 4a, such an effect was not observed. Although termination factor rho induces the synthesis of some minor RNA products and the amount of the original RNA products is somewhat reduced, the sizes of the three major RNA species remain unaltered. From this result we infer that the strong, rho-dependent termination signal(s) are not located on fragment *Hae* III-A and hence must be located between the 3'-terminal end of this fragment and promoter $G_{0.08}$.

Restriction fragment *Hap* II-A is 1530 nucleotides long and encompasses the complete gene IV as well as a small part of the intergenic region (20). In the absence of rho, this fragment directs the synthesis of only a single RNA species which is ini-

tiated at promoter $A_{0.36}$ and reaches a length of about 1450 nucleotides (2, Fig. 3g). If fragment Hap II-A also encompasses the strong, rho-dependent, termination signal(s), then the length of the rho-induced product will be determined by promoter $A_{0.36}$ and the position of these termination signal(s) on fragment Hap II-A. However, transcription of fragment Hap II-A in the presence of rho, gives rise to the synthesis of many prematurely terminated products the two largest of which, denoted 1 and 2, are predominant (Fig. 3h). Since fragment Bam. Hind-A covers fragment Hap II-A completely (Fig. 1), it is expected that rho-terminated RNA products which are formed on fragment Hap II-A have their equivalents among the rho-terminated products formed on fragment Bam. Hind-A. Comparing Fig. 3e with Fig. 3h, shows that under the direction of fragment Bam. Hind-A two RNA products are formed which are similar in length to the rho-terminated products 1 and 2 formed upon transcription of fragment Hap II-A. It is therefore probable that these two RNA products originate in both cases from promoter $A_{0.36}$. The sizes of the latter products (1300 and 1450 nucleotides long respectively) indicate that the termination signals responsible for these two rho-dependent termination events are located at the ultimate 3'-end of fragment Hap II-A. Furthermore, the difference in length between the two RNA species suggest that there are in fact two termination signals operative which are located about 150 nucleotides apart. Similar observations have been made for E.coli and for phage ØX-174. On these two templates consecutive, rho-dependent termination signals separated by a stretch of about 180 base pairs have been detected in the gene coding for tyrosine $tRNA_{su3}$ and in the boundary of the genes J and F respectively (17, 21, 22).

The existence of termination sites at the 3'-terminal end of fragment Hap II-A should also reduce the sizes of the RNA species initiated at the promoters $A_{0.56}$ and $A_{0.51}$ (Fig. 1). One might, therefore, expect that in the presence of rho the two transcripts initiated at these promoters are truncated to lengths of 2700 (i.e. 22S) and 2300 nucleotides (i.e. 20S) respectively. That rho-induced RNA products of such a size are indeed formed upon transcription of fragment Bam. Hind-A (Fig. 3e, bands 3 and 4) and up-

on transcription of intact RF (Fig. 2a, f, bands 3 and 4) supports our conclusion concerning the position of the rho-dependent termination site. It is, however, remarkable that with RF-I as template the synthesis of the 20S RNA species is relatively strong (cf. Fig. 2a, f and Fig. 3e). Since this pronounced synthesis disappears as soon as a single cleavage site is introduced in the RF-I molecule, by cleavage of the RF molecule with endo R. Hind II for example (data not shown), this effect may reside in a conformational change of the template DNA.

Effect of rho on the central termination site

Previously we reported that, under in vitro conditions and in the absence of rho, only a single transcription termination signal located between the genes VIII and III is operative on the M13 genome (19). Following our observation that termination at this central termination signal is only about 90% effective, it has been postulated that readthrough products contribute to the expression of genes which are located distal to this signal (18, 4). To estimate whether rho affects this readthrough phenomenon, we have studied the transcription of restriction fragment Hap II-B_2 (Fig. 1).

Transcription of fragment Hap II-B_2 results in the formation of a major RNA product of about 370 nucleotides long which is terminated at the central termination signal (23,16). Additionally, a minor RNA product of about 700 nucleotides long is generated. Previously we suggested that the synthesis of the latter transcript is a result of leakage of RNA polymerase molecules through the central termination site (18). Hence, synthesis of the minor transcript can be used to examine the effect of rho on the efficiency of the termination signal. Comparison of Fig. 4c with Fig. 4d shows that the addition of rho to the transcription system decreases the synthesis of the minor RNA product. From these data the termination efficiency at the central termination signal was calculated to increase from 90% up to 95% suggesting that the central termination signal is under partial control of termination factor rho. However, this conclusion cannot be made as long as the secondary, possibly nonspecific, ef-

fects of rho which have been noticed to occur at numerous sites on the M13 RF molecule (cf. Figs. 3a, e, h and 4a) are not excluded.

DISCUSSION

The transcription studies presented were carried out in order to locate possible rho-dependent termination signals on the genome of the filamentous phage M13. However, it is very difficult to distinguish between true and artificial rho-mediated termination events and therefore we have focussed our attention on the localisation of strong termination signals which are operative under moderate rho concentrations. The preliminary results of these studies suggest that, among numerous sites where weak termination events occur, there exists a unique region on the phage M13 genome where rho-mediated termination events occur very efficiently. In studies in which the intact M13 RF molecule and the restriction fragments Bam. Hind-A, Hae III-A and Hap II-A were used as templates, this region was found to be located within the boundary of gene IV and the intergenic region. Support for the presence of a termination signal within this part of the phage genome was recently provided by *in vivo* transcription studies (6). In these studies phage-specific mRNA species from M13 RF harbouring mini-cells of *E.coli* were analysed by means of gel electrophoresis. It was demonstrated that among the *in vivo* products no RNA species were detectable equivalent to the *in vitro* and in the absence of rho generated 23S and 26S RNA species. On the other hand, all of the smaller *in vitro* synthesized RNA species (8S up to 19S) were present among the mini-cell RNA products. These results suggest,therefore, that under *in vivo* conditions, the synthesis of only the pppA-initiated RNA chains is terminated before the rho-independent central termination site is reached. Since this *in vivo* transcription pattern shows a striking resemblance to the transcription pattern which is obtained under *in vitro* conditions and in the presence of rho, it is tempting to speculate that the rho-sensitive termination signal distal to gene IV is also operative *in vivo*. If this is true, then the genome of the filamentous phages must be regarded as

consisting of (at least) two different transcriptional units which are defined by a rho-independent and a rho-dependent termination site. According to the position of these two strong termination signals, the smallest of the two units covers the genes II, X, V, VII and VIII whereas the other unit covers the genes III, VI, I and IV. Consequently, the smaller unit contains all G-promoters and the larger unit all A-promoters and the gene III promoter.

According to our nucleotide sequence studies (23), the central termination signal on the phage M13 genome exhibits several features which are characteristic for rho-independent termination signals. Due to these features the 3'-ends of RNA species terminating at the central termination signal contain an ultimate series of eight uridenylate residues, a sequence block rich in G-C residues and, additionally, these RNA regions have the intrinsic property of forming a hairpin-like structure. Although some of these properties also have been reported for 3'-ends of RNA species terminated at rho-dependent termination signals (24) recent studies by Küpper et al. (22) have indicated that the 3'-terminal regions of transcripts terminated at completely rho-dependent termination signals may contain other features. According to their data, the rho-dependent termination signals present in the tyrosine $tRNA_{su3}$ gene of E.coli are characterized by the sequence CAATCAA and the presence of two adjacent DNA sequences which are extremely rich in G-C residues. Following these observations we were interested to determine whether a DNA sequence with these properties or, alternatively, with the properties described for partial rho-dependent termination signals, are also present within that region of the M13 genome which covers the boundary of gene IV and the intergenic region. An examination of the relevant nucleotide sequence data indicated, however, that a structure with features identical to those described for partially or completely rho-dependent termination signals does not exist within this particular part of the phage M13 genome (3; Van Wezenbeek and Schoenmakers, unpublished results). One must, therefore, conclude that the rho-dependent termination signals which are operative on this template DNA have other characteristics. Alternative but very attractive candidates for the latter

signals occur on the M13 genome within the intergenic region at sites which are located at 10 and 135 nucleotides distal to gene IV respectively. These two sites form part of extensively base-paired hairpin structures and share several features with the completely rho-dependent termination signals described by Küpper et al. For example, they contain an A-T rich sequence preceding a sequence which extremely rich in G-C base pairs. Additionally, a cluster of G-C base pairs is located immediately upstream of the A-T rich sequence. However, neither of these two presumed termination sites contains the characteristic sequence CAATCAA described for the tyrosine tRNA gene. Instead, the equivalent regions on the phage M13 genome contain the sequences CATTAAG and CTCTAAA respectively. Whether or not the latter sequences form part of the rho-dependent termination signals which are operative on the M13 genome awaits further investigation, but it is rather intriguing that their positions fully account for the results described in this paper.

REFERENCES

1. Konings, R.N.H. and Schoenmakers, J.G.G. (1978). In: *Single-Stranded DNA Phages*. (D.T. Denhardt, D.S. Ray and D. Dressler, eds.) Cold Spring Harbor Laboratory, in press.
2. Edens, L., van Wezenbeek, P., Konings, R.N.H. and Schoenmakers, J.G.G. (1976). *Eur. J. Biochem.*, 70, 577-587.
3. Schaller, H. and Takanami, M. (1978). In: *Single-Stranded DNA Phages*. (D.T. Denhardt, D.S. Ray and D. Dressler, eds.) Cold Spring Harbor Laboratory, in press.
4. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). *Virology*, XX, 000-000.
5. Jacob, E., Jaenisch, R. and Hofschneider, P.H. (1973). *Eur. J. Biochem.*, 32, 432- 443.
6. Smits, M.A., Simons, G., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). *Biochem. Biophys. Acta*, XX, 000-000.
7. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1973). *Mol. Biol. Rep.*, 1, 41-45.
8. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 547-558.
9. Van den Hondel, C.A. and Schoenmakers, J.G.G. (1976). *J. Virol.*, 18, 1024-1039.
10. Van den Hondel, C.A., Weyers, A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Eur. J. Biochem.*, 53, 559-567.
11. Roberts, J.W. (1969). *Nature*, 224, 1168-1174.
12. Darlix, J.L. (1975). *Eur. J. Biochem.*, 51, 369-376.
13. Konings, R.N.H., Hulsebos, T. and van den Hondel, C.A. (1975). *J. Virol.*, 15, 570-584.
14. Konings, R.N.H. and Bloemendal, H. (1969). *Eur. J. Biochem.*, 7, 165-173.
15. Seeburg, P.H. and Schaller, H. (1975). *J. Mol. Biol.*, 92, 261-277.
16. Sugimoto, K., Sugisaki, H., Okamoto, T. and Takanami, M. (1977). *J. Mol. Biol.*, 110, 487-507.
17. Axelrod, N. (1976). *J. Mol. Biol.*, 108, 753-770.
18. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1978). *Virology*, 86, 354-367.

19. Edens, L., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Nucleic Acid Res.*, 2, 1811-1820.
20. Van den Hondel, C.A., Konings, R.N.H. and Schoenmakers, J.G.G. (1975). *Virology*, 67, 487-497.
21. Axelrod, N. (1976). *J. Mol. Biol.*, 108, 771-779.
22. Küpper, H., Sekiya, T., Rosenberg, M., Egan, J. and Landy, A. (1978). *Nature*, 272, 423-428.
23. Edens, L., (1978). *Ph. D. Thesis*, University of Nijmegen.
24. Roberts, J.W. (1976). In: *RNA Polymerase* (R. Losick and M. Chamberlin eds.) Cold Spring Harbor Laboratory, pp. 247-271.

SUMMARY

Bacteriophage M13 belongs to the group of F-specific filamentous phages (Ff-phages) which make use of the bacterium Escherichia coli for their multiplication. However, infection with these phages is restricted to bacteria that contain F-pili. These F-pili are long filaments which are extruded from bacteria and which play an important role in both the bacterial conjugation event and the adsorption of Ff-phages. In the initial phase of the infection process, the Ff-phages attach selectively to the tip of the F-pilus after which they are transported to the cell surface, probably as a result of retraction of the F-pilus. Immediately after penetration, the DNA molecule present in the phage particle is replicated and the phage genome is expressed. These processes ultimately lead to the synthesis of progeny phages which are extruded by the bacterium. Since the release of filamentous phages is not accompanied by lysis of the host-cell, phage production, which is initiated about twenty minutes after infection, continues for hours and can give rise to the synthesis of a few thousands of progeny phage particles.

In the first stage of the DNA replication process, the single-stranded DNA molecule of the filamentous phage is converted into a double-stranded, circularly closed replicative form. This latter, so called RF molecule, plays a key role in the generation of new, viral DNA strands and the synthesis of phage-encoded proteins. Synthesis of the various phage-specific polypeptides is accomplished by transcription of the RF molecule with E.coli RNA polymerase and a subsequent translation of the mRNA molecules formed. However, the proteins generated are not made in equimolar amounts but instead important quantitative differences arise. To explain the origin of these quantitative differences, it is attractive to speculate that either the number of mRNA molecules available is regulated or that particular mRNA species are translated preferentially.

In the experiments described in this thesis an attempt was made to elucidate the nature of the mechanisms which regulate the expression of the bacteriophage M13 genome. For this purpose the transcription mechanism of the M13 RF molecule was analysed

with the aid of *in vitro* transcription and translation systems. In the first chapter, describing the experimental results (Chapter III) the various mRNA molecules formed upon transcription of M13 RF with *E.coli* RNA polymerase holoenzyme were characterized. Analysis of the *in vitro* synthesized RNA products by means of gel electrophoresis indicated that at least seven discrete RNA species are formed which are initiated either with guanosine triphosphate (pppG) or with adenosine triphosphate (pppA). The pppG-initiated RNA chains were found to range in size from about 360 nucleotides (8S) up to about 2000 nucleotides (19S). The sizes of the pppA-initiated RNA chains varied from about 3500 nucleotides (23S) up to about 5000 nucleotides (26S). Since all of these RNA products hybridise to denatured RF and not to the viral DNA, it was concluded that in the *in vitro* transcription process only the complementary DNA strand is copied. This observation is in good agreement with the results obtained in hybridisation experiments reported previously in which phage-specific RNA products from M13 infected *E.coli* cells were analysed. The hybridisation studies, therefore, indicate that *in vivo* and *in vitro* RNA synthesis proceeds in the same direction along the phage genome. Moreover, the data presented in Chapter III indicate that under transcription conditions used, there is only a single transcription termination signal operative on the M13 genome. By using various restriction fragments in both transcription and translation experiments, it could be demonstrated that this so called 'central' termination signal is located distal to the termination codon of gene VIII at a distance of not more than 60 base pairs.

Since RNA synthesis proceeds in only one direction along the M13 RF molecule and terminates at only one specific site, the seven RNA species which are generated in the transcription process are most likely the result of the presence of seven RNA-initiation sites on the phage genome. However, the transcription studies on isolated restriction fragments described in Chapter IV show that at least eight such 'promoter' sites are operative on this template molecule. At five of these promoters RNA synthesis is initiated by the incorporation of pppG. Initiation at the remaining three promoters takes place by the incorporation of

pppA. By measuring the length of the RNA fragments synthesized and by assuming that upon transcription of restriction fragments RNA synthesis is terminated at the end of these fragments, the various promoters were precisely located on the phage genome. The experiments demonstrated that the five G-promoters are clustered within a distance of one-third of the genome length proximal to the central termination site for transcription. The three A-promoters, however, were found at greater distances from this site. Furthermore, labeling studies with (γ-^{32}P) and (γ-^{32}P) ATP showed important differences between the RNA-initiating capacities of the various promoters. That is, on the basis of the amounts of radioactivity incorporated, it appears that for both G-promoters and A-promoters the maximum differences in their initiation capacities was approximately one order of magnitude.

In order to locate the five G-promoters precisely on the genetic map of the M13 genome, restriction fragments containing the central transcription termination signal and, in addition, one or more promoters were isolated. By using these fragments in both transcription and coupled transcription-translation experiments, it was demonstrated (Chapter V) that each of the genes VIII, V, X and II had its 'own' promoter which was located immediately in front of the relevant gene. Moreover, it was shown that besides the promoter in front of gene X, another G-promoter occurs within gene II. However, up to now the function of the latter promoter remains obscure. The length and the codogenic capacity of the pppG-initiated transcripts indicated that all G-promoters give rise to the formation of an RNA chain, the synthesis of which is not terminated before the central termination site for transcription is reached. This implies that the 3'-terminal ends of the various RNA chains overlap and thus increase the transcription frequencies of those genes which are located proximal to the central termination signal. Such a 'cascade-like' transcription mechanism, therefore, offers a qualitative explanation for the high amounts of gene V and gene VIII proteins which are synthesized during expression of the RF molecule. However, a preferential synthesis of the latter two proteins was also observed upon the translation of isolated, polycistronic, mRNA molecules. This demonstrates that the expression of the bacterio-

phage M13 genome is not only regulated on the level of transcription, but also on the level of translation. A good example of the latter type of regulation is the expression of gene VII. Though this gene is transcribed very frequently, its protein product has never been detected.

In view of the importance of the central termination signal for RNA synthesis in the expression of the M13 genome, Chapter VI, is devoted to a discussion of the structure of this termination signal and possible mechanisms causing termination of transcription. Information concerning the structure of the termination signal was obtained by nucleotide sequence analysis of the 3'-terminal part of an in vitro synthesized RNA species. According to these data, the termination signal consists of a series of eight A-T base pairs preceded by an inverted DNA sequence repeat rich in G-C base pairs. The consequence of the presence of this specific nucleotide sequence on the template DNA is that the RNA transcribed from the termination signal is terminated in a region rich in uridenylate residues. Additionally, the 3'-terminal end of the RNA product has the intrinsic property of forming a tight, hairpin-like, structure. Since similar features have recently been described for transcripts terminated at signals occurring on the genome of bacteriophage lambda and in the tryptophan operon of E.coli, it is very likely that the hairpin structure as well as the A-T and G-C rich sequence blocks are directly involved in the transcription termination event.

According to our in vitro experiments, termination of RNA synthesis at the central termination signal is not 100% effective. About 10% of the RNA polymerase molecules leak through the termination signal thereby increasing the transcription frequencies of the genes located downstream of this signal. Since that part of the M13 genome which is located immediately distal to the central termination signal contains only promoters which are relatively weak, it is expected that the 'readthrough' RNA products influence the expression of the genes III, VI and I significantly. However, the experiments described in Chapter VII do not support this hypothesis. According to the latter data the expression of the genes III, VI and I is primarily dependent upon promoters which are located distal to the central termination

signal. This result compares well with experiments in which gene I, gene VI and, in contrast to previous observations, gene III also were found to be equiped with their 'own' promoters. How the latter data can be reconciled with the previously observed polarity among these three genes remains unresolved.

The discrepancy which exists regarding the sizes of the *in vivo* and *in vitro* synthesized, phage-specific, mRNA molecules, can be explained by assuming the presence of additional transcription termination signals on the M13 genome. To test whether rho-dependent termination signals are involved, the effect of rho on the transcription pattern of the M13 RF molecule was analysed. As described in Chapter VIII, the addition of purified rho protein to the *in vitro* transcription system specifically affects the synthesis of the larger RNA species. This observation together with the sizes of various, rho-terminated, RNA products strongly suggests that the effect of rho on the transcription pattern of the M13 RF molecule is due to the presence of rho-dependent termination signals in the boundary of gene IV and the so called 'intergenic region'. As a result of the presence of these additional termination signals, the synthesis of the three pppA-initiated RNA chains is, in the presence of rho, terminated before the rho-independent, central termination signal has been reached. So, on the basis of these data, the M13 genome consists of two transcriptional units the smallest of which contains all G-promoters and the other all A-promoters and the gene III promoter. If similar units exist *in vivo*, then the transcription frequencies of the various M13 genes are possibly regulated by the cascade transcription mechanism as well as by the rho-dependent termination signal.

Bacteriofaag M13 behoort tot de groep van de F-specifieke filamenteuze fagen (Ff-fagen) die voor hun vermenigvuldiging afhankelijk zijn van de darmbacterie Escherichia coli. Infectie is echter alleen mogelijk indien de gastheercel in het bezit is van F-pili. Deze F-pili zijn lange, flagelachtige structuren die naast hun rol in de bacteriële conjugatie ook een duidelijke functie vervullen in de adsorptie van de Ff-fagen. In de eerste fase van het infectieproces binden de Ff-fagen zich namelijk specifiek aan het uiteinde van de F-pilus waarna, vermoedelijk door middel van retractie van de F-pilus, de fagen naar het cel oppervlak getransporteerd worden. Direct na het binnendringen van het faagpartikel komen er in de gastheercel een aantal ingewikkelde processen op gang die er voor zorgen dat het virale DNA molecule gerepliceerd wordt en de virale genen tot expressie komen. Deze processen leiden tenslotte tot de synthese van nieuwe faagpartikels die door de bacteriecel in het medium worden afgescheiden. Omdat het vrijkomen van filamenteuze fagen niet gepaard gaat met lysis van de gastheer, kan de faagproductie, die ca. 20 minuten na infectie start, uren doorgaan en aanleiding geven tot de vorming van een paar duizend nakomelıng fagen.

In de initiële fase van het infectie proces wordt het enkelstrengs DNA-molecule van de filamentueze faag omgezet in een dubbelstrengs, circulair gesloten replicatieve vorm. Dit laatste zogenaamde RF-molecule, vervult een sleutelrol in de aanmaak van nieuwe virale DNA moleculen en in de synthese van de diverse virale eiwitten. Synthese van de faag-specifieke eiwitten komt tot stand doordat het RF-molecule getranscribeerd wordt door het E.coli RNA-polymerase waarna de gevormde mRNA-moleculen vertaald worden. De in dit proces gevormde eiwitten worden echter niet in equimolaire hoeveelheden gesynthetiseerd maar er ontstaan aanzienlijke kwantitatieve verschillen tussen de diverse genproducten. Een aannemelijke verklaring voor het ontstaan van deze verschillen is dat er tijdens de expressie van het Ff-genoom een constitutief regulatiemechanisme werkzaam is dat zorgt voor een preferente transcriptie van bepaalde virale genen en/of voor een preferente translatie van bepaalde transcriptieproducten.

In de experimenten die in dit proefschrift beschreven staan hebben we geprobeerd de aard te doorgronden van de mechanismen die de expressie van het bacteriofaag M13 genoom reguleren. Hiervoor hebben we het transcriptiemechanisme van het M13 RF-molecule bestudeerd met behulp van een *in vitro* transcriptie- en een *in vitro* translatiesysteem. In hoofdstuk III hebben we door middel van gelelectroforese de diverse mRNA moleculen gekarakteriseerd die gevormd worden als gevolg van transcriptie van het M13 RF-molecule met *E.coli* RNA-polymerase holoenzyme. Uit de verkregen resultaten blijkt dat er minstens zeven discrete RNA klassen ontstaan die geïnitieerd worden met guanosinetrifosfaat (pppG) of met adenosinetrifosfaat (pppA). De lengtes van de met pppG geïnitieerde RNA klassen varieren van ca. 360 nucleotides (8S) tot ca. 2000 nucleotides (19S) en de lengtes van de met pppA geïnitieerde RNA-klassen van ca. 3500 nucleotides (23S) tot ca. 5000 nucleotides (26S). Uit het feit dat de diverse RNA-producten alleen hybridiseren aan gedenatureerd RF en niet aan viraal DNA, blijkt dat in het *in vitro* transcriptieproces alleen de complementaire DNA streng gecopieerd wordt. Deze waarneming is in overeenstemming met de resultaten van eerder beschreven hybridisatie experimenten waarin faagspecifieke RNA-producten uit M13-geïnfecteerde *E.coli* cellen werden geanalyseerd. Op grond van deze resultaten moeten we dus concluderen dat *in vivo* én *in vitro* de transcriptiereactie in één en dezelfde richting langs het faaggenoom verloopt. Uit de in hoofdstuk III beschreven experimenten blijkt tevens dat *in vitro* slechts een enkel transcriptietermineringssignaal werkzaam is op het M13-genoom. Met behulp van transcriptie- en translatiestudies hebben we aangetoond dat dit zogenaamde "centrale" RNA-termineringssignaal achter het stopcodon van gen VIII is gelegen op een afstand van maximaal 60 basenparen.

Aangezien de RNA-synthese slechts in één richting langs het M13 RF-molecule verloopt en bovendien slechts op één specifieke plaats op het faaggenoom termineert, is het zeer waarschijnlijk dat de zeven RNA-klassen die in het transcriptieproces ontstaan een direct gevolg zijn van de aanwezigheid van zeven RNA-initiatieplaatsen op het faaggenoom. Door middel van in hoofdstuk IV beschreven transcriptiestudies aan geïsoleerde restrictiefrag-

menten is echter gebleken dat er sprake is van minstens acht van dergelijke "promoter sites". De initiatie van de RNA-synthese aan deze promotors vindt plaats door de incorporatie van pppG (G-promotors) of door de incorporatie van pppA (A-promotors). Door aan te nemen dat terminering van de RNA-synthese plaatsvindt aan het eind van de desbetreffende restrictiefragmenten, zijn we in staat geweest om, via de lengtes van de gesynthetiseerde RNA-fragmenten, zowel de G-promotors als de A-promotors nauwkeurig op het faaggenoom te lokaliseren. Uit deze experimenten is gebleken dat er vijf G-promotors voorkomen in dat gedeelte van het faaggenoom wat direkt voor het transcriptieterminerings signaal gelegen is. De drie overige promotors (A-promotors) bleken daarentegen op grotere afstanden van dit signaal te liggen. Studies waarin de diverse RNA-producten gemerkt werden met (γ^{32}P)GTP of (γ^{32}P)ATP hebben bovendien laten zien dat er belangrijke verschillen bestaan tussen de RNA-initiërende capaciteiten van de diverse promotors. Op grond van geïncorporeerde hoeveelheden radioactiviteit is namelijk gebleken dat, zowel bij de G-promotors als bij de A-promotors, onderlinge capaciteitsverschillen voorkomen die maximaal een factor 10 bedragen.

Teneinde de exacte plaats van de vijf G-promotors op de genetische kaart van het M13-genoom te bepalen, hebben we gebruik gemaakt van een serie restrictiefragmenten die, behalve het centrale termineringssignaal voor de RNA-synthese ook één of meer promotors bevatten. Door middel van transcriptie- en translatie-experimenten met deze fragmenten bewijzen we in hoofdstuk V dat de genen VIII, V, X en II over een "eigen" promotor beschikken, die direct voor het desbetreffende gen is gelegen. Bovendien blijkt er sprake van een G-promotor welke, naar analogie met de promotor voor gen X, gelegen is in gen II maar waaraan echter geen duidelijk genproduct kan worden toegeschreven. Uit de lengtes en de codogene eigenschappen van de diverse RNA-ketens kon bovendien worden afgeleid dat alle G-promotors aanleiding geven tot de vorming van een RNA-molecule waarvan de synthese niet eerder getermineerd wordt voordat het centrale termineringssignaal bereikt is. Dit houdt in, dat de 3'-terminale uiteinden van de diverse RNA-moleculen elkaar overlappen waardoor de transcriptiefrequentie van die genen die direct voor het centrale termi-

neringssignaal liggen sterk verhoogd worden. Dit zogenaamde "cascade" transcriptiemechanisme biedt op grond van deze eigenschappen dus een kwalitatieve verklaring voor de hoge synthese snelheid van de eiwitten die gecodeerd worden door de genen V en VIII. Aangezien echter andere in dit hoofdstuk beschreven experimenten duidelijk hebben laten zien dat beide eiwitten ook preferent gesynthetiseerd worden bij de vertaling van geïsoleerde, polycistronische mRNA-ketens, moeten we concluderen dat regulatie van de genexpressie niet alleen op transcriptieniveau maar ook op translatieniveau plaatsvindt. Een duidelijk voorbeeld van dit laatste wordt gevormd door gen VII. Het eiwitproduct van dit gen is namelijk nog nooit gedetecteerd, niettegenstaande het feit dat het gen gelegen is tussen gen V en gen VIII en derhalve zeer frequent getranscribeerd wordt.

Gezien het belang van het centrale transcriptietermineringssignaal in de expressie van het M13-genoom, is hoofdstuk VI gewijd aan de structuur van dit signaal en aan mogelijke mechanismen die leiden tot terminering van de RNA-synthese. Gegevens over de structuur van het centrale termineringssignaal zijn verkregen door het bepalen van de nucleotiden sequentie van het 3'-terminale uiteinde van een in vitro gesynthetiseerd RNA-product. Aan de hand van deze gegevens bleek dat het centrale termineringssignaal voor de RNA-synthese bestaat uit een zogenaamde "inverted repeat" van twee G-C rijke gebieden die gevolgd wordt door een reeks van acht A-T basenparen. Gezien het feit dat dergelijke eigenschappen recentelijk ook beschreven zijn voor transcriptietermineringssignalen die voorkomen op het genoom van faag lambda en in het tryptofaan operon van E.coli, is het zeer waarschijnlijk dat zowel de "inverted repeat" structuur als de reeks A-T basenparen direct betrokken zijn bij het termineren van de RNA-synthese.

Volgens onze transcriptie-experimenten verloopt de terminering van de RNA-synthese door het centrale termineringssignaal niet 100% effectief. Ongeveer 10% van de RNA-polymerase moleculen passeren het termineringssignaal ongehinderd en verhogen hierdoor de transcriptiefrequentie van de genen die direct achter het termineringssignaal gelegen zijn. Gezien het feit dat in dit gedeelte van het M13-genoom slechts relatief zwakke promotors gelegen zijn, zouden de in het lekproces gevormde "readthrough" RNA-

producten een duidelijke bijdrage moeten leveren aan de expressie van met name de genen III, VI en I. Echter, de in hoofdstuk VII beschreven translatie-experimenten steunen deze veronderstelling niet maar suggereren daarentegen dat de expressie van deze drie genen in de eerste plaats afhankelijk is van promotors die direct voor de genen III, VI en I zijn gelegen. In hoeverre deze laatste gegevens te rijmen zijn met de resultaten van genetische complementatiestudies die op grond van polariteitseffecten doen vermoeden dat de drie genen tot expressie komen via een enkele, polycistronische mRNA keten, wordt niet duidelijk.

De discrepantie die bestaat ten aanzien van de lengtes van in vivo en in vitro gesynthetiseerde faagspecifieke mRNA-moleculen suggereren dat in vivo additionele transcriptieterminerings-signalen werkzaam zijn op het M13-genoom. Om na te gaan of het hier termineringssignalen betreft die afhankelijk zijn van de transcriptietermineringsfactor rho, hebben we het effect van dit eiwit op het transcriptiepatroon van het M13 RF-molecule geanalyseerd. Zoals beschreven in hoofdstuk VIII beïnvloedt het toevoegen van gezuiverd rho aan het in vitro transcriptiesysteem de synthese van met name de grotere RNA-moleculen. Lengtemetingen aan rho-getermineerde RNA producten maken het waarschijnlijk dat dit een gevolg is van de aanwezigheid van een rho-afhankelijk termineringssignaal in de zogenaamde "intergenic region". Als gevolg van dit additionele termineringssignaal wordt in de aanwezigheid van rho de synthese van de diverse pppA-geïnitieerde RNA-ketens getermineerd voordat het centrale, rho-onafhankelijke termineringssignaal is bereikt. In aanwezigheid van rho bestaat het M13 genoom dus uit twee transcriptionele units waarvan de kleinste unit alle G-promotors omvat en de grootste unit alle A-promotors en de gen III promotor. Indien dergelijke units ook in vivo zouden bestaan, zou dit inhouden dat de transcriptiefrequenties van de diverse M13 genen op tweeërlei wijzen gereguleerd kunnen worden: enerzijds door middel van het cascade transcriptiemechanisme en anderzijds door middel van het rho-afhankelijke termineringssignaal.

CURRICULUM VITAE

De schrijver van dit proefschrift werd in 1949 in Groningen geboren. In 1969 behaalde hij in deze zelfde stad zijn eindditploma HBS-B waarna hij biologie ging studeren aan de plaatselijke Rijksuniversiteit. Na het behalen van het kandidaatsexamen B4 en het voltooien van hoofdvakken op de afdelingen Moleculaire Genetica (Dr. G. Venema) en Microbiologie (Prof. H. Veldkamp), werd de studie in januari 1972 afgerond met het doctoraal examen. Tot september 1973 werd de militaire dienstplicht vervuld, waarna per september van datzelfde jaar een aanstelling volgde als wetenschappelijk medewerker bij het Laboratorium voor Moleculaire Biologie van de Katholieke Universiteit te Nijmegen. In de daarop volgende vier jaren werd het experimentele werk verricht dat in dit proefschrift beschreven is. In deze periode trouwde hij met Astrid Heymeriks en werd Jef geboren. In zijn huidige functie is hij werkzaam bij Unilever Research Vlaardingen waar hij sinds 1 januari, 1978 deel uitmaakt van de groep "Genetic Engineering".

STELLINGEN

1. De bewering van Shen en Hearst dat er een duidelijke correlatie bestaat tussen de positie van promotors op het genoom van bacteriofaag fd en het voorkomen van secondaire structuren in het virale DNA is volkomen speculatief.

   Shen, C.K.J. en Hearst, J.E. (1976) Proc. Natl. Acad. Sci. U.S.A., 73, 2649-2653.

2. De aandacht die in sequentie-studies wordt besteed aan het voorkomen van palindroom structuren in DNA- en RNA-moleculen berust meer op een veronderstelde dan op een bewezen biologische rol van deze structuren.

3. Handbook No. 23 van het International Biological Programme bevat diverse weinig deugdelijke methodes voor de vaststelling van microbiële activiteit in zoet water.

   IBP Handboek No. 23
   Y.I. Sorokin en H. Kadota, eds.
   Blackwell Scientific Publications.

4. Wijzigingen van de oppervlaktespanning in het water-lucht-grensvlak van eiwitoplossingen zoals gemeten door Katona et al., hoeven niet noodzakelijkerwijs het gevolg te zijn van conformatie veranderingen.

   Katona, E. Neumann, A.W. en Moscarello, M.A. (1978). BBA 534, 275-284.

5. Het feit dat het vermogen tot tumorvorming (nopaline synthese) bij tabaksplanten verloren gaat bij generatieve vermeerdering, beperkt het gebruik van het Ti-plasmide als vector in de transformatie van planteceilen tot die planten die vegetatief vermeerderd kunnen worden.

   Braun, A.C. en Wood, N.H. (1976) Proc. Natl. Acad. Sci. U.S.A., 73, 496-500.

6. Het tot expressie komen van plasmiden die in Bacillus subtilis worden geïntroduceerd door middel van transformatie, vereist renaturatie of recombinatie van deze DNA-moleculen.

Canosi, U., Morelli, G. en Trautner, T.A. (1978). MGG, in druk.

7. De vaak geciteerde opvatting van Johannsen dat Clostridium botulinum type E een grondbacterie is, wordt alleen gesteund door de experimenten van Johannsen zelf.

Johannsen, A. (1965). J. Appl. Bact. 28, 90-94.

8. De conclusie van Chattopadhyay et al. dat het aantal virus-specifieke DNA-sequenties in cellulair DNA niet toeneemt als gevolg van infectie met muizen-leukemie virus is aanvechtbaar.

Chattophadhyay, S.K., Rowe, W.P. en Levine, A.S. (1976). Proc. Natl. Acad. Sci. U.S.A., 73, 4095-4099.

9. Gezien de research activiteiten in de ons omringende landen zal het verbieden van recombinant-DNA-onderzoek Nederland niet vrijwaren van de eventuele macro-gevolgen van dit soort onderzoekingen.

10. Uit het feit dat de Tweede Kamer het beslist noodzakelijk acht dat ook vrouwen zitting nemen in de Commissie belast met het toezicht op recombinant-DNA-onderzoek, zou men kunnen opmaken dat nog niet iedereen op de hoogte is van het feit dat dit soort experimenten slechts met prokaryoten en eenvoudige eukaryoten wordt uitgevoerd.

Verslag Tweede Kamer der Staten-Generaal, zitting 1977-1978, 14438, nr. 4.

L. Edens, Nijmegen, 15 december, 1978.